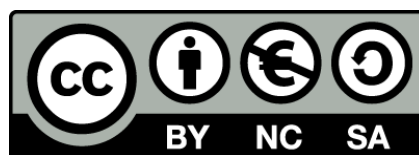




UNIVERSITAT DE
BARCELONA

The role of the histone variant macroH2A in the regulation of chromatin architecture

David Corujo Garcia



Aquesta tesi doctoral està subjecta a la llicència **Reconeixement- NoComercial – CompartirIgual 4.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento - NoComercial – CompartirIgual 4.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution-NonCommercial-ShareAlike 4.0. Spain License.**

Universitat de Barcelona

Facultat de Biologia

Programa de Doctorat en Genètica

The role of the histone variant macroH2A in the regulation of chromatin architecture

Memòria presentada per

David Corujo García

per optar al grau de doctor per la Universitat de Barcelona

Tesi realitzada a

l'Institut de Recerca contra la Leucèmia Josep Carreras

Director

Tutora

Doctorand

Marcus
Buschbeck

Montserrat
Corominas Guiu

David
Corujo García

Barcelona, setembre 2019

Als meus avis

Acknowledgements

El doctorat és una experiència complexa, plena d'alts-i-baixos, il·lusió, frustració, creativitat, planificació i improvisació. Durant aquests anys sento que he crescut enormement tant a nivell científic com a nivell personal, adquirint coneixements específics però, sobretot i encara més important, habilitats i maneres de pensar i actuar transversals que penso seguir desenvolupant i sento que em seran útils per tota la vida. Per suposat, durant aquesta etapa d'aprenentatge i creixement he comptat amb moltes persones al meu voltant dins i fora del laboratori a les quals els hi dec un molt sincer agraïment.

Quiero dar las gracias a Marcus por haberme dado la oportunidad de hacer el doctorado en el laboratorio. Todavía recuerdo el primer día que vine al IMPPC como estudiante, desde el primer momento me ofreciste la posibilidad de implicarme en los proyectos y siempre me he sentido muy respetado en el trabajo, con la posibilidad de discutir cualquier idea contigo. Además he tenido la posibilidad de formarme en diferentes aspectos de la carrera investigadora por lo que te estoy muy agradecido.

Julien, has sido mi mentor desde mis primeros pasos en el laboratorio y en los primeros años del doctorado. Has invertido mucho tiempo y esfuerzo en mi formación, incluso en los momentos más difíciles, y te estaré por siempre agradecido. He aprendido muchísimo de ti, desde protocolos hasta diseñar experimentos y pensar de manera crítica, buscar soluciones y hacer valer mis ideas. Somos dos personas muy diferentes pero recuerdo con mucho cariño el tiempo que pasamos juntos y echo de menos tus consejos, ideas y bromas. Aún a día de hoy, si tengo problemas, muchas veces me quedo pensando: WWJD? (what would Ju do?).

Vane, moltes gràcies per tot el suport i paciència durant aquests anys. Estic molt content d'haver-te tingut al costat (literalment!) durant aquest temps, no només per feina si no també personalment. He après molt de tu, tant tècniques com organització i els últims mesos treballant colze a colze han estat genials. Tenim maneres d'entendre la feina i la vida semblants, cada cop m'he sentit més còmode al teu costat i saber que puc compartir les vivències amb tu ho ha fet tot més fàcil.

Sarah, we started our PhDs together and have gone through similar ups and downs over the course of the years, sharing scientific discussions, countless coffee breaks, meetings and trips. Month by month I gained a friend in my colleague. You have always been by my side in difficult periods trying to cheer me up. In many aspects, we have very

different attitudes but have always found a way to fit together and I have learned many things from you. Thank you for everything.

Roberto, muchas gracias por todas las ideas y trabajo que has aportado a los proyectos, estoy muy contento de poder trabajar contigo y tengo muchas ganas de seguir compartiendo proyectos (y ensayos) contigo. Además, me has ayudado mucho en mis “aventuras” en el mundo de la programación, sin ti no habría podido progresar tanto. Y por supuesto te llevas el premio a la mejor cultura musical del laboratorio!

Iva and Marguerite-Marie, thanks for all the interesting discussions we have shared and your participation in the project. Ainhoa, muchas gracias por tu dedicación y ganas de aprender. Michael, Raquel and Jeannine, thanks for making the lab a supportive environment and always be willing to provide help and input. Anna, encara trobo a faltar el teu suport en el bàndol “nacional” del lab, compartint bromes de l’APM i la competència, sempre em vas ajudar quan ho vaig necessitar, moltes gràcies. I also thank past members of the laboratory, Neus and Mel, you were always very helpful and supportive when I was just a “kid” starting out.

Moltes gràcies als col·laboradors que hem tingut al Campus Can Ruti. Carolina Armengol, Laura Royo i Juan Carrilo, gràcies pel vostre suport i aportacions en el projecte. Arce Garcia i Josep Manyé, sense la vostra participació no hauria estat possible desenvolupar el projecte d’experimentació animal i totes les idees que ens ha proporcionat. Arce, sobretot, moltes gràcies per la teva implicació en la feina conjunta que hem fet, he après molt treballant amb tu.

Many thanks to our collaborators in Munich at the group of Andreas G. Ladurner, specially Marek Kozlowski, Imke Mandemaker and Charlotte Blessing, for your hard work. I hope we can continue to share ideas and work together in the future.

A part dels membres del laboratori i col·laboradors, molta gent al campus ha fet d’aquests anys una gran experiència. Mar, muchas gracias por ayudarme siempre en cualquier cosa que he necesitado, ha sido genial ser labs vecinos en el IMPPC y recibir tus “piropos” ;) Patri, muchas gracias también a ti por ayudarme desde la “cocina” cuando solo era un “pipiolo”, me hace mucho ilusión que después de tanto tiempo volvamos a estar trabajando cerca. Gràcies també a molts d’altres investigadors i tècnics dels campus amb els que he compartir dinars, cafés i que sempre heu estat disposats a donar-me un cop de mà. Muchas gracias al “board game geek” club de Can Ruti, Roberto, José, Jorge,

Bernat, Yaiza, Edu, Marguerite-Marie, Llorenç... por tantas horas de diversión y desconexión del trabajo compartiendo uno de mis hobbies favoritos.

Moltes gràcies als meus amics i amigues dins i fora del món de la investigació per tots els sopars, cerveses i escapades que han omplert aquest període.

Moltes gràcies a la meva família, avis, tiets, cosins, pares, que m'han fet costat en tot moment, s'han interessat per la meva feina i han sigut un punt de suport i referència essencial en la meva vida. Als meus avis, perquè sempre m'han transmès un profund amor pel coneixement i la cultura i gràcies als seus esforços avui puc gaudir de l'oportunitat d'escriure aquesta tesi. Als meus pares, que sempre m'han donat suport i han valorat totes les decisions que he pres, m'han educat en el valor de l'esforç i el pensament crític i m'han proporcionat les eines i habilitats per enfrontar-me als reptes de la meva carrera i gaudir-ne. Gràcies al meu germà Guillem, per estar sempre al meu costat, no seria el mateix sense haver compartit la vida amb tu. Per acabar, moltes gràcies a la Susanna, per estar sempre amb mi en els bons moments i en els dolents, no concebo aquest viatge sense haver-te tingut al costat. Hem viscut junts una època plena de canvis i reptes, i sóc molt feliç d'haver-los superat junts. En certa manera tanco una etapa, i estic molt il·lusionat de començar-ne encara una de nova, amb tu.

Table of Contents

ABBREVIATIONS	1
INTRODUCTION	7
1. EPIGENETICS AND CHROMATIN	9
1.1. THE CONCEPT OF EPIGENETICS	9
1.2. CHROMATIN IS THE TEMPLATE FOR EPIGENETIC REGULATION	10
1.3. MECHANISMS OF CHROMATIN REGULATION	12
1.3.1. <i>DNA methylation</i>	12
1.3.2. <i>Post-translational modifications of histones</i>	13
1.3.3. <i>Histone variants</i>	16
2. MACROH2A HISTONE VARIANTS	19
2.1. THE MACROH2A FAMILY OF HISTONE VARIANTS	19
2.2. MOLECULAR AND PHYSIOLOGICAL FUNCTIONS OF MACROH2A	22
2.2.1. <i>Nucleosomal stability and structure</i>	22
2.2.2. <i>Genomic localization and transcriptional regulation</i>	22
2.2.3. <i>DNA repair of double-strand breaks</i>	23
2.2.4. <i>Positive factor in differentiation and senescence</i>	24
2.2.5. <i>Context-dependent tumor suppression</i>	25
2.2.6. <i>Impact on NAD⁺ metabolism</i>	26
3. HETEROCHROMATIN AND THE THREE DIMENSIONAL ORGANIZATION OF THE GENOME.....	27
3.1. HETEROCHROMATIN	27
3.1.1. <i>Constitutive heterochromatin and repetitive elements</i>	28
3.1.2. <i>Facultative heterochromatin and polycomb</i>	31
3.2. BASIC PRINCIPLES OF 3D GENOME ORGANIZATION	34
3.2.1. <i>Chromosome territories</i>	35
3.2.2. <i>A/B compartments</i>	36
3.2.3. <i>Domains associated with the nuclear lamina and the nucleolus</i>	37
3.2.4. <i>Topologically associating domains (TADs) and the loop extrusion model</i>	40
3.2.5. <i>Local interactions between genes and regulatory elements</i>	43
AIMS AND OBJECTIVES	47
4. AIMS AND OBJECTIVES	49
RESULTS.....	51
5. PREAMBLE	53
6. RESULTS I: MACROH2A MAINTAINS NUCLEAR ORGANIZATION AND HETEROCHROMATIN ARCHITECTURE.....	55
6.1. MACROH2A IS ESSENTIAL FOR MAINTAINING NUCLEAR ORGANIZATION	55
6.2. A FRACTION OF MACROH2A IS ASSOCIATED WITH H3K9ME3 AND REPETITIVE ELEMENTS	58
6.3. MACROH2A MAINTAINS THE ARCHITECTURE OF HETEROCHROMATIC REPETITIVE ELEMENTS	61
6.4. MACROH2A PROMOTES THE ATTACHMENT OF REPETITIVE ELEMENTS TO LAMIN B1 FILAMENTS	64
6.5. CONTRIBUTIONS	68
7. RESULTS II: MACROH2A CAN LIMIT CHROMATIN PLASTICITY THROUGH TWO DISTINCT MECHANISMS	71
7.1. MACROH2A PROTEINS SHOW DIFFERENT TISSUE EXPRESSION PATTERNS AND CONSERVED STRUCTURAL FEATURES	71
7.2. MACROH2A1.1, BUT NOT MACROH2A2, BINDS ADP-RIBOSE AND INHIBITS PARP1	76
7.3. MACROH2A HISTONES SUPPRESS CHROMATIN DYNAMICS UPON DNA DAMAGE THROUGH TWO DISTINCT MECHANISMS....	78
7.4. THE LINKER REGION OF MACROH2A LIMITS CHROMATIN EXPANSION UPON DNA DAMAGE IN LIVING CELLS.....	83
7.5. THE MACROH2A LINKER REGION IS ESSENTIAL AND SUFFICIENT FOR STABILIZING HETEROCHROMATIN ARCHITECTURE	85
7.6. CONTRIBUTIONS	87

8. RESULTS III: MACROH2A REPRESSES THE EXPRESSION OF DKK1 AND ATTENUATES ITS TRANSCRIPTIONAL ACTIVATION BY TNFA.....	89
8.1. HEPG2 CELLS LACKING MACROH2A SHOW AN INCREASED COLONY FORMATION AND MIGRATION CAPACITY, BUT ARE UNAFFECTED IN CELL GROWTH.....	89
8.2. TUMORS DERIVED FROM MACROH2A-LACKING CELLS DO NOT HAVE INCREASED GROWTH OR HIGHER LEVELS OF MAJOR MALIGNANCY MARKERS	89
8.3. TRANSCRIPTOMIC PROFILING OF HEPG2-DERIVED TUMORS SHOWS MAJOR GENE EXPRESSION CHANGES THAT DEPEND ON MACROH2A	92
8.4. GENES ALTERED IN THE ABSENCE OF MACROH2A ARE INVOLVED IN CELLULAR ADHESION, DEVELOPMENT, HYPOXIA AND INFLAMMATION	94
8.5. <i>DKK1</i> IS UPREGULATED IN HEPG2-DERIVED TUMORS LACKING MACROH2A AND HEPATOBLASTOMA TUMORS	98
8.6. THE ABSENCE OF MACROH2A RESULTS IN AN UPREGULATION OF <i>DKK1</i> EXPRESSION IN RESPONSE TO AN INFLAMMATORY STIMULUS	99
8.7. MACROH2A1.1 AND MACROH2A2 CAN SUPPRESS <i>DKK1</i> ACTIVATION IN RESPONSE TO TNFA	101
8.8. ABSENCE OF MACROH2A ALTERS THE 3D STRUCTURE OF THE <i>DKK</i> LOCUS AND PERMITS A STRONGER ACTIVATION OF DISTAL REGULATORY ELEMENTS UPON TNF α TREATMENT	102
8.9. CONTRIBUTIONS	103
DISCUSSION	107
9. DISCUSSION I	109
9.1. MACROH2A IS REQUIRED FOR MAINTAINING HETEROCHROMATIN ARCHITECTURE AND ATTACHMENT TO THE NUCLEAR LAMINA	109
9.2. POSSIBLE MECHANISMS OF MACROH2A-LAMIN B1 INTERACTION	109
9.3. MACROH2A IS NOT ESSENTIAL FOR THE SILENCING OF REPETITIVE ELEMENTS.....	110
9.4. MACROH2A MAINTAINS NUCLEAR AND NUCLEOLAR INTEGRITY	111
9.5. THE FUNCTION OF MACROH2A IN HIGHER-ORDER STRUCTURES MAY NOT BE LIMITED TO CONSTITUTIVE HETEROCHROMATIN	112
10. DISCUSSION II.....	115
10.1. MACROH2A HISTONE VARIANTS SHOW DIVERSITY IN STRUCTURE, EXPRESSION AND FUNCTION	115
10.2. MACROH2A1.1 CAN REGULATE PARP1-DEPENDENT PROCESSES.....	115
10.3. POSSIBLE MECHANISMS OF STABILIZATION OF CHROMATIN STRUCTURES BY THE MACROH2A LINKER.....	116
10.4. POTENTIAL REGULATION OF THE MACROH2A LINKER BY PTMS.....	119
11. DISCUSSION III.....	121
11.1. MACROH2AS ARE CONTEXT- AND ISOFORM- DEPENDENT TUMOR SUPPRESSORS THAT IMPACT ON TRANSCRIPTIONAL REGULATION	121
11.2. MACROH2A REPRESSES <i>DKK1</i> , A GENE UPREGULATED IN HEPATOBLASTOMA AND LIVER CANCER	122
11.3. MACROH2A CAN REGULATE THE TRANSCRIPTIONAL RESPONSE TO HYPOXIA AND INFLAMMATORY SIGNALLING	124
11.4. THE MAINTENANCE OF EPIGENETIC AND TRANSCRIPTIONAL STATES BY MACROH2A MIGHT BE MEDIATED BY THE STABILIZATION OF 3D CHROMATIN STRUCTURES	126
CONCLUSIONS.....	129
12. CONCLUSIONS	131
MATERIALS AND METHODS	133
13. MATERIALS AND METHODS.....	135
13.1. ANTIBODIES	135
13.2. PLASMIDS	135
13.3. CELL CULTURE	136
13.3.1. <i>Standard cell culture conditions</i>	136
13.3.2. <i>Colony formation assay</i>	136
13.3.3. <i>Wound healing assay</i>	137
13.3.4. <i>XTT proliferation assay</i>	137
13.3.5. <i>TNFα treatment</i>	137
13.4. GENE TRANSDUCTION.....	137
13.4.1. <i>Transfection</i>	137

13.4.2. Retroviral infection	138
13.5. PROTEIN EXPRESSION ANALYSIS	138
13.5.1. Protein extraction from cultured cells	138
13.5.2. Precipitation of extracellular proteins	139
13.5.3. Protein extraction from mouse tissues	139
13.5.4. Western blot (SDS-PAGE)	139
13.6. RNA EXPRESSION ANALYSIS	140
13.6.1. RNA isolation	140
13.6.2. cDNA synthesis by reverse transcription	140
13.6.3. Quantitative real-time PCR (RT-qPCR)	140
13.6.4. In situ mRNA hybridization	141
13.6.5. RNA-Seq and differential gene expression analysis	141
13.6.6. Gene ontology enrichment analysis	142
13.6.7. Gene Set Enrichment Analysis (GSEA)	142
13.7. CHROMATIN IMMUNOPRECIPITATION (CHIP)	142
13.7.1. ChIP-qPCR	142
13.7.2. ChIP-Seq and association analysis	144
13.8. UMI-4C	145
13.9. MICROSCOPY	146
13.9.1. Immunohistochemistry	146
13.9.2. Immunofluorescence staining	147
13.9.3. Fluorescence In Situ Hybridization (FISH) and immuno-FISH	147
13.9.4. Proximity ligation assay (PLA)	148
13.9.5. Fluorescence recovery after photobleaching (FRAP)	148
13.9.6. Transmission electron microscopy	148
13.9.7. Imaging of immunofluorescence staining and FISH	148
13.9.8. Imaging of DNA damage induced by laser micro-irradiation	149
13.9.9. Image analysis	149
13.10. FORMALDEHYDE-ASSISTED ISOLATION OF REGULATORY ELEMENTS	150
13.11. PROTEIN EXPRESSION AND PURIFICATION	150
13.12. PROTEIN CRYSTALLIZATION AND STRUCTURE DETERMINATION	151
13.13. ISOTHERMAL TITRATION CALORIMETRY (ITC)	152
13.14. PARP1 <i>IN VITRO</i> ACTIVITY ASSAY	152
13.15. XENOGRAFT	153
13.16. STATISTICS AND PLOTS	153
13.17. ANTIBODIES TABLE	154
13.18. PRIMERS TABLE	155
BIBLIOGRAPHY	157
14. BIBLIOGRAPHY	159

Abbreviations

Abbreviations

ADPr	ADP-ribose
ATRX	Alpha Thalassemia/Mental Retardation Syndrome X-Linked
BER	Base excision repair
BRCA1	Breast Cancer 1
CBP	CREB-binding protein
CBX	Chromobox protein
CDK8	Cyclin Dependent Kinase 8
cHC	Constitutive heterochromatin
ChIP	Chromatin immunoprecipitation
cLAD	Constitutive LAD
cPRC1	Canonical polycomb repressive complex 1
CTCF	CCCTC-Binding Factor
DamID	DNA adenine methyltransferase identification
DAXX	Death Domain Associated Protein
DEG	Differentially expressed gene
DKK1	Dickkopf 1
DNMT	DNA methyltransferase
EED	Embryonic ectoderm development
EM	Electron microscopy
EMT	Epithelial to mesenchymal transition
ERBB2	Erb-B2 Receptor Tyrosine Kinase 2
ESC	Embryonic stem cell
EZH1/2	Enhancer of zeste 1/2
f/vLAD	Facultative/variable LAD
FISH	Fluorescence <i>in situ</i> hybridization
FRAP	Fluorescence recovery after photobleaching
GLP	G9a-like protein
GO	Gene Ontology
GSEA	Gene set enrichment analysis

HAT	Histone acetyltransferase
HB	Hepatoblastoma
HDAC	Histone deacetylase
HiC	High-throughput chromosome conformation capture
HIF	Hypoxia inducible factor
HP1	Heterochromatin protein 1
IF	Immunofluorescence
INM	Inner nuclear membrane
ITC	Isothermal titration calorimetry
KDM	Histone lysine demethylase
KMT	Histone lysine methyltransferase
LAD	Lamina-associated domain
LAP2β	Lamina-associated Polypeptide 2 isoform beta
LBR	Lamin B receptor
LINE	Long interspersed nuclear element
LMNB1	Lamin B1
Low comp.	Low complexity
LOX	Lysil oxidase
LTR	Long terminal repeat
Major/minor sat.	Major/minor satellite
MBD	Methyl-CpG binding domain protein
MEF	Mouse embryonic fibroblast
mH2A	macroH2A
mH2A1	macroH2A1.1 and macroH2A1.2
mH2A1.1	macroH2A1.1
mH2A1.2	macroH2A1.2
mH2A2	macroH2A2
MMP	Matrix metalloproteinase
NAD	Nucleolus-associated domain
NAP-1	Nucleosome assembly protein 1
ncPRC1	Non-conincal polycomb repressive complex

NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NK	Natural killer
NPC	Nuclear pore complex
NPM1	Nucleophosmin 1
o/n	Overnight
ONM	Outer nuclear membrane
PARP1	Poly(ADP-ribose) polymerase 1
PC	Principal component
PcG	Polycomb group of proteins
PCGF	Polycomb Group Ring Finger
PFA	Paraformaldehyde
PLA	Proximity ligation assay
PRC	Polycomb repressive complex
PTM	Post-translational modification
RC	Rolling circle
rDNA	Ribosomal DNA
rRNA	Ribosomal RNA
RT	Room temperature
RYBP	RING1 And YY1 Binding Protein
scRNA	Small conditional RNA
SETDB1/2	SET Domain Bifurcated 1/2
shRNA	Short-hairpin RNA
Simple rep.	Simple repeat
SINE	Short interspersed nuclear element
snRNA	Small nuclear RNA
srpRNA	Signal recognition particle RNA
SUV39H1/2	Suppressor Of Variegation 3-9 Homolog 1/2
SUZ12	Supressor of zeste 12
TAD	Topologically associating domains
TEM	Transmission electron microscopy
TET	Ten-eleven translocation proteins

TNFα	Tumor necrosis factor alpha
tRNA	Transfer RNA
Xi	Inactive X chromosome
ZEB1	Zinc Finger E-Box Binding Homeobox 1

Introduction

1. Epigenetics and chromatin

1.1. The concept of epigenetics

The coinage of the term “epigenetics” is usually credited to the British biologist Conrad H. Waddington who, in 1942, defined epigenetics as follows: “the branch of biology which studies the causal interactions between genes and their products, which bring the phenotype into being” (Waddington, 1942). Waddington was an embryologist and thought of epigenetics in the context of the developmental process of an organism. To this day, the early development of complex multicellular organisms is still widely used as the prime example of epigenetic mechanisms at work. During embryonic development, a single totipotent cell gives rise to a whole organism composed of a myriad of cells that share an identical genomic sequence but have very different gene expression profiles and cellular functions. Waddington conceived this process of cellular differentiation as a marble rolling down a slope with grooves that sequentially bifurcate (**Figure 1**). The marble at the top representing a totipotent cell would roll through this “epigenetic landscape” (Waddington, 1957), progressively being “canalized” through the grooves until reaching the lowest stable points, that is, a differentiated “end state”.

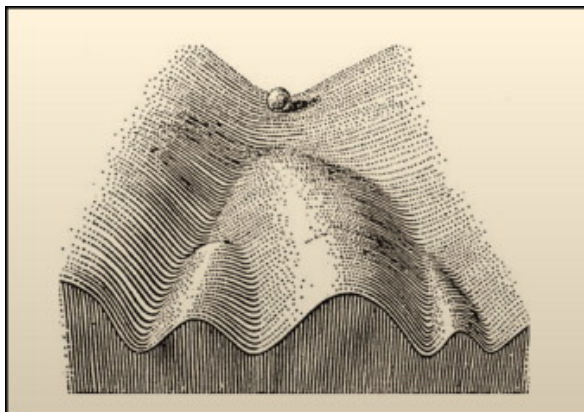


Figure 1. Waddington's epigenetic landscape. The classical depiction of the “epigenetic landscape” as conceived by Waddington, where the rolling of a marble down the grooves of a slope represents a cell during development taking different paths to achieve a differentiated cell state. Taken from Goldberg et al., 2007.

At the time when Waddington coined the term “epigenetics”, the physical nature of the genes and the molecular mechanisms regulating their function was still unknown, but he still conceptualized epigenetics as a bridge to explain the complex and dynamic interactions between genotype, environment and phenotype. This concept remains associated to the field of epigenetic studies to this day but, more precisely, the term “epigenetics” is commonly used to describe the mechanisms of genetic regulation involving factors other than the sequence of the DNA molecule.

1.2. Chromatin is the template for epigenetic regulation

In the eukaryotic cell nucleus, DNA is tightly wrapped around histone protein complexes, forming a dynamic structure known as chromatin (Luger, 2006). The nucleosome particle is the basic repeating unit of the chromatin fiber. The nucleosome consists of a histone octamer core around which 147bp of DNA are wrapped about 1.7 times (Luger et al., 1997; McGinty and Tan, 2015). The protein core of the nucleosome is composed of the so-called core histones: H2A, H2B, H3 and H4 (**Figure 2**). Core histones are highly conserved proteins that consist of an unstructured N-terminus, a globular histone fold core and a mobile C-terminal tail. In particular, two units each of H3 and H4 form a tetramer around which bind two dimers of H2A-H2B. Additionally, the H1 linker histone binds the nucleosome and the extra-nucleosomal DNA, stabilizing chromatin structure and allowing higher levels of compaction (Hergeth and Schneider, 2015).

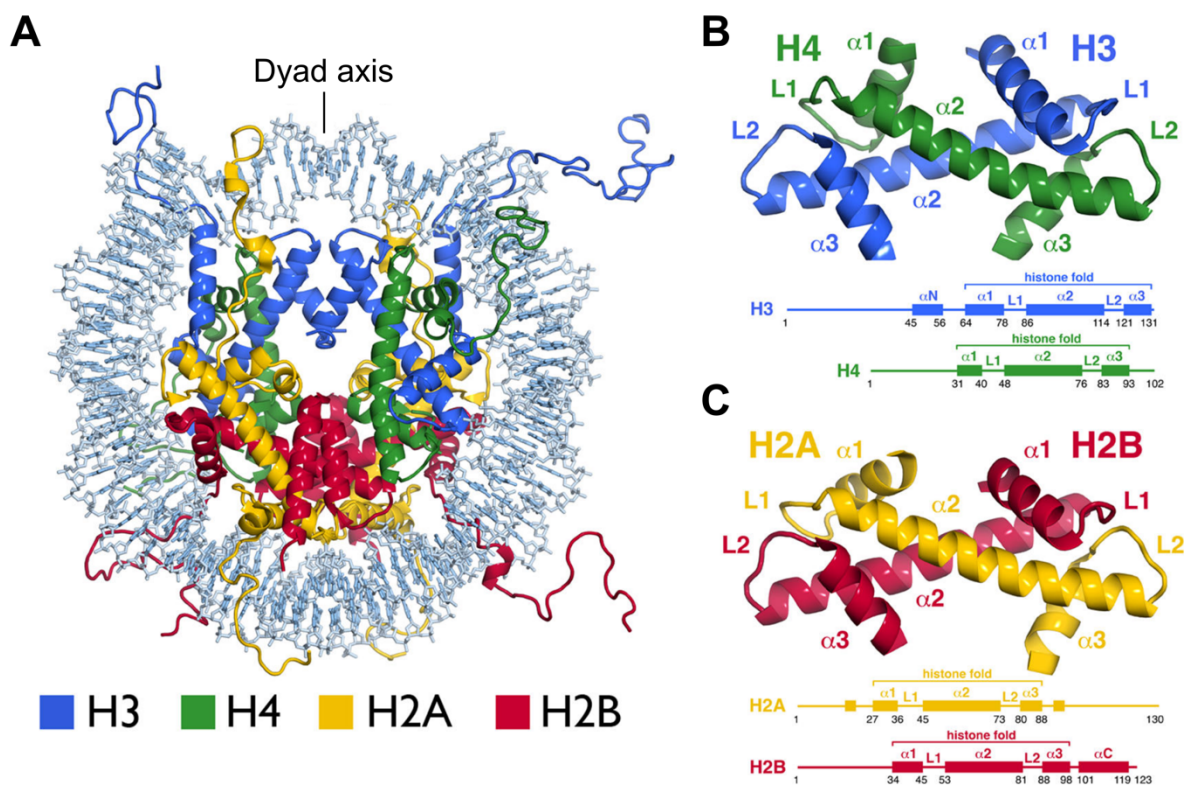


Figure 2. The structure of the nucleosome core particle. (A) Structure of the nucleosome core particle based on PDB ID 1KX5 showing the compact histone octamer with DNA wrapped around it. Histones and DNA are represented in cartoon and sticks style, respectively. Histones are colored as indicated in the legend. (B, C) Structure of the core histone dimers H3-H4 and H2A-H2B. The top panel shows a representation of each dimer in cartoon style while the bottom panels show a linear schematic representation of each histone. In both cases, the main secondary structures are indicated (α refers to α -helices and L refers to loops). Figure adapted from McGinty and Tan, 2015.

The commonly accepted model of chromatin folding presents different levels of hierarchical compaction forming fibers with increasing size, allowing for the packaging of thousands of megabases of DNA in the small nuclear volume which measures only a few micrometers in diameter (**Figure 3**). The nucleosomes are separated by a stretch of linker DNA ranging from 10 to 90 bp forming the so-called “beads-on-a-string” 11nm fiber, named after its appearance in electron microscopy images (Olins et al., 1974). The second structural level of chromatin organization after the 11 nm fiber involves the H1 linker histone allowing the condensation of nucleosomal arrays into proposed solenoid or zig-zag type structures forming the 30nm-fiber (Grigoryev and Woodcock, 2012). However, the precise structural nature of some of the higher-order levels of chromatin compaction, and the very existence and characteristics of the 30nm-fiber in the nucleus of living cells, remain questioned to this day (Maeshima et al., 2019). Recent studies using super-resolution microscopy have revealed an arrangement in the form of heterogeneous groups of nucleosomes termed “clutches” (Fussner et al., 2012; Ricci et al., 2015). In any case, it is generally considered that nucleosome positioning and the level of chromatin compaction regulate the accessibility of the transcriptional machinery and repair factors, among others.

Thus, we have to consider that in eukaryotes the genetic information encoded in the DNA molecule is contained in chromatin and, as such, changes happening in chromatin can potentially impact all processes of DNA function like transcription, replication and recombination, among others. Far from being solely a packaging system, chromatin is a highly dynamic complex that can be altered at many levels without modification of the underlying DNA sequence, effectively forming the basis and template of epigenetic regulation.

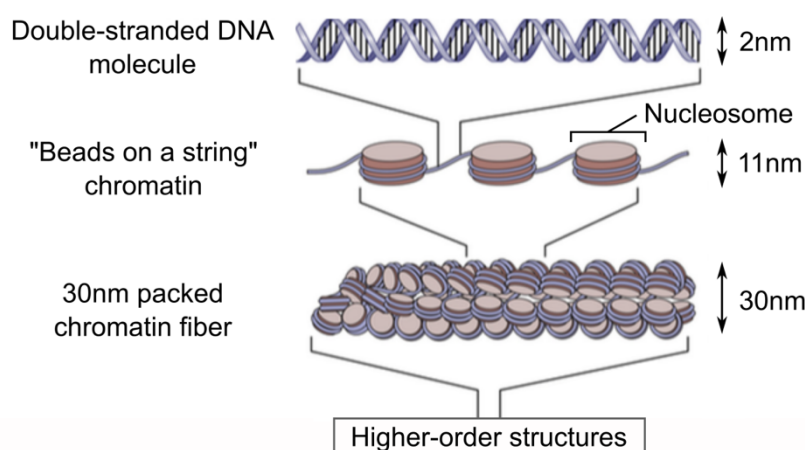


Figure 3. The classic model of chromatin folding in the nucleus. The double-stranded DNA molecule is wrapped around nucleosomes separated by a stretch of linker DNA. This “beads on a string” chromatin is packed into a 30nm chromatin fiber. These fibers are further compacted into higher-order structures that have not been characterized in detail. Figure adapted from Jansen and Verstrepen, 2011.

1.3. Mechanisms of chromatin regulation

The mechanisms by which the properties of the chromatin are modified are commonly referred to as “epigenetic modifications”. These modifications can consist of the actual covalent modification of either the DNA molecule or histone proteins, like in the case of DNA methylation or the acetylation of histones. The balance between enzymatic “writer” and “eraser” factors determines the presence of a covalent mark that can be interpreted by “reader” factors and effector proteins (**Figure 4**).

Other mechanisms, however, consist of non-covalent modifications that involve the incorporation of specialized histone variants in the nucleosome, changes in chromatin accessibility by chromatin remodelers, or the interaction with non-coding RNA molecules. The ensemble of epigenetic modifications on the chromatin of a given cell is commonly referred to as the “epigenome”. Below, I will individually discuss three of the best characterized epigenetic modifications to date regarding their molecular mechanisms as well as their known functions in physiology and pathological conditions.

1.3.1. DNA methylation

Initially discovered in 1975 (Holliday and Pugh, 1975; Riggs, 1975), the methylation of DNA is probably the best characterized epigenetic modification to date. In mammals, nearly all methylation events of the DNA molecule occur within a CpG dinucleotide context by modification of a cytosine base to a 5-methylcytosine (5meC). Regions of the genome with a high density of CpGs are referred to as CpG islands, 1 to 2 kb regions that frequently contain gene promoters (Illingworth and Bird, 2009). Active promoters are usually characterized by unmethylated CpG islands, while methylated CpG islands correlate with transcriptional repression (Jones, 2012). On the other hand, methylation in the gene body correlates with transcriptional elongation and prevents spurious transcription initiation (Neri et al., 2017).

DNA methyltransferases (DNMTs) form a conserved family of enzymes responsible for establishing and maintaining DNA methylation patterns (Lyko, 2018). The process of demethylation can either passively happen as a consequence of dilution during DNA replication or actively through the step-wise oxidation of 5meC by the ten-eleven translocation proteins (TET) followed by the base excision repair (BER) that restores an unmethylated cytosine (Rasmussen and Helin, 2016).

DNA methylation has a prominent and essential role in development and differentiation (Smith and Meissner, 2013). Its alteration in cancer cells has been extensively studied and currently some cancer types are treated with agents that inhibit DNMTs, for instance acute myeloid leukemia and myelodysplastic syndromes (Jones et al., 2016).

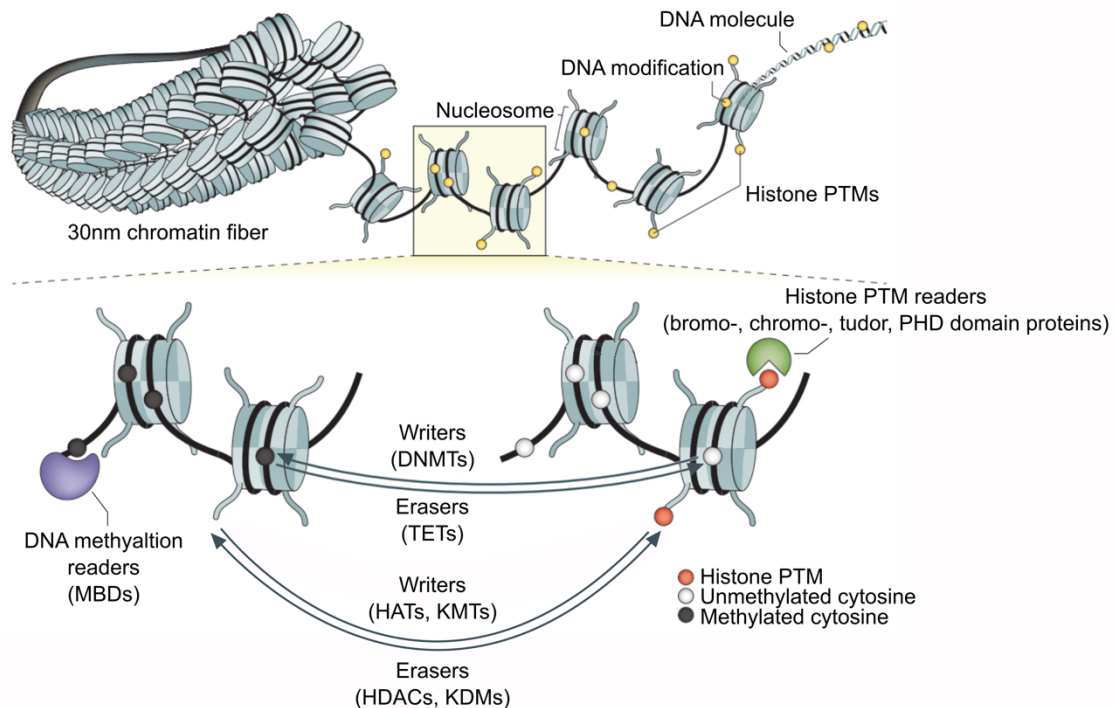


Figure 4. Chromatin regulation by covalent modifications. Several mechanisms of epigenetic regulation involve covalent modification of the DNA molecule or histone proteins. The presence of a particular mark is controlled by the action of “writer” and “eraser” factors that deposit and remove it, respectively. These covalent modifications can be bound by specialized domains in “reader” factors that can trigger direct effects or recruit other effector proteins. Some examples of protein families involved in each process are given between parenthesis: DNMTs, DNA methyltransferases; TETs, ten-eleven translocation; HATs, histone acetyltransferases; KMTs, histone lysine methyltransferases; HDACs, histone deacetylases, KDMs, histone lysine demethylases; MBDs, methyl-CpG binding domain proteins). Figure adapted from Jones and Baylin, 2016.

1.3.2. Post-translational modifications of histones

Post-translational modifications (PTMs) are an essential and highly dynamic mechanism for the regulation of protein function and the transduction and integration of signals in the cell (Deribe et al., 2010). This is remarkably relevant for histone proteins as their covalent modifications contribute to the molecular basis of epigenetic regulation and cellular memory.

Introduction

The highly basic and unstructured N-termini of all four histones and the C-termini of H2A protrude out of the core structure of the nucleosome, while the C-termini of the other histones is buried in the nucleosome particle core (**Figure 2**). The first described histone PTM was acetylation and its association to actively transcribed genes was quickly established (Allfrey et al., 1964; Pogo et al., 1966). To date, more than 100 different histone PTMs have been reported, including acetylation, methylation, phosphorylation, ubiquitination, SUMOylation, citrullination and ADP-ribosylation, among others (Zhao and Garcia, 2015). Given their accessible position, it is not surprising that most histone PTMs happen at the N-terminus. However, PTMs in the globular domain of histones have also been described to affect nucleosomal structure and transcription (Lawrence et al., 2016). Usually, histone PTMs are abbreviated by combining the name of the histone, the one-letter code and position of the modified residue and the shorthand for the covalently linked group, potentially adding a number to indicate the number of such groups. For example, the trimethylation of H3 at lysine 27 would be abbreviated as H3K27me3.

In general, histone PTMs are thought to participate in chromatin regulation by two major mechanisms. First, the covalent addition of chemical groups to histones can directly alter the net charge or other physico-chemical properties of histone proteins, affecting DNA-nucleosome or internucleosomal interactions and the packaging of chromatin. For example, the acetylation of lysine residues at histone tails neutralizes positive charges and effectively weakens the charge-dependent interactions between histones and nucleosomal DNA, increasing its accessibility (Hong et al., 1993). Similarly, ADP-ribosylation of histones imparts negative charge creating a relaxed chromatin conformation through repulsion (Messner and Hottiger, 2011). Second, histone PTMs can serve as a binding module to recruit “reader” proteins that can specifically recognize PTMs via specialized domains such as bromo-, chromo- and PHD-domains (Musselman et al., 2012). These readers can then carry out effects on chromatin and/or further recruit other effector proteins. Importantly, these two mechanisms are not exclusive: in addition to directly affecting nucleosomal structure, histone acetylation can be recognized by bromodomain-containing proteins that, alone or as part of larger complexes, can regulate the transcriptional machinery, induce chromatin remodeling or provoke further histone PTMs (Fujisawa and Filippakopoulos, 2017).

Different histone modifications frequently co-occur in the same nucleosome. Mutually exclusive relations exist if the modification takes place on the same residue or if there is compartmentalization or incompatibility between the writer factors. Still, histone

PTMs present a combinatorial nature with a high degree of cross-talk, where the presence of a PTM can avoid, antagonize, facilitate or be a requirement of another one. This led to the proposal of the “histone code” hypothesis by which PTMs on histone tails are interpreted sequentially or in combination by reader proteins or complexes to provoke different downstream effects, essentially functioning as an epigenetic “code” on chromatin (Strahl and Allis, 2000). More recent approaches to this idea have used the computational integration of genome-wide profiling of histone PTMs and other chromatin components to functionally annotate the whole genome in different “chromatin states” (Ernst and Kellis, 2010).

Although they are found both in genes and non-coding regions of the genome and participate in multiple chromatin processes, histone PTMs have been extensively studied in the context of gene regulation and classified as either “active” or “inactive” depending on their relation with transcription. For example, active marks include H3K4me3 and H3K36me3, which are prominently associated with the promoter and gene body of actively transcribed genes, respectively. On the other hand, H3K27me3 and H2AK119Ub are enriched in the chromatin of transcriptionally repressed genes. Interestingly, active and inactive marks are not always mutually exclusive: H3K4me3 and H3K27me3 are found together in “bivalent” promoters of developmental genes in embryonic stem cells, granting a state that is silent but poised for rapid activation (Bernstein et al., 2006; Harikumar and Meshorer, 2015a). Histone PTMs are also found in non-coding distal regulatory elements. Most notably, H3K4me is found at enhancer elements while H3K27Ac distinguishes active from inactive enhancers (Creyghton et al., 2010).

Histone PTMs are an essential mechanism that is linked to many chromatin function processes beyond transcription like DNA replication and DNA repair (van Attikum and Gasser, 2009; Lukas et al., 2011). They are highly dynamic modifications that modulate gene expression programs during development and cellular differentiation. Moreover, mutations and deregulation in the factors that orchestrate this epigenetic control are frequent in cancer and crucial in the control of the expression of oncogenes and tumor suppressor genes (Jones and Baylin, 2007; Shen and Laird, 2013). This has sparked great interest in the development of “epigenetic drugs” for the treatment of cancer targeting histone modifications “writer” and “eraser” enzymes as well as “reader” modules. For example, different inhibitors of histone de-acetylases (HDACi) are currently used in the treatment of lymphomas, while inhibitors of domains binding to acetylated histones

are being clinically tested for the treatment of hematological malignancies and solid tumors (Jones et al., 2016).

1.3.3. Histone variants

The bulk of the histone pool in the cell is composed of "canonical" or replication-coupled histones, which are encoded in tandem array gene clusters and synthesized and deposited into chromatin in a replication-dependent manner. A smaller fraction of the histone pool is constituted by histone variants, that diverge to different extents in their primary sequence from their replication-coupled counterparts (Maze et al., 2014). Histone variants are generally expressed from single-copy genes and deposited into chromatin independently of replication by specialized machinery (Buschbeck and Hake, 2017; Gurard-Levin et al., 2014). Histone variants are dynamically regulated both in their expression level and their genomic deposition. The replacement of replication-coupled core histones by histone variants provides chromatin with specific characteristics and can thus influence all functions occurring on the chromatin template including transcription and DNA repair (Biterge and Schneider, 2014; Buschbeck and Hake, 2017). Similarly to histone modifications, this influence can occur through different mechanisms that include: alteration of the biophysical properties of the nucleosome (Bönisch and Hake, 2012), promoting the deposition of certain histone modifications or recruitment of specific interactors (Buschbeck and Hake, 2017).

In humans, several variant forms have been described for the H2A, H2B and H3 core histones (**Figure 5**), some of them are "universal" and highly conserved among all eukaryotes while others have evolved specifically in higher eukaryote lineages (Talbert and Henikoff, 2010). Additionally, different variants of the H1 linker histones have been identified (Hergeth and Schneider, 2015). H2A variants represent the largest and most diverse family of histones. In human somatic cells, eight variants of H2A have been identified: H2A.X, H2A.Z.1, H2A.Z.2.1, H2A.Z.2.2, H2A.Bbd, macroH2A1.1, macroH2A1.2 and macroH2A2. As for H3, six variants are described: H3.3, CENP-A, H3.1T, H3.5, H3.X (also known as H3.Y.2) and H3.Y (also known as H3.Y.1). Additionally, several germ cell-specific histones have been identified that function in spermatogenic differentiation and paternal genome activation after fertilization (Hoghoughi et al., 2017; Huynh et al., 2016; El Kennani et al., 2017; Shinagawa et al., 2014).

Analogous to histone PTMs, histone variants are enriched and function in different genomic contexts. Some of them have very specialized roles in particular chromosomal regions. CENP-A, for example, is exclusively present at centromeres and is essential for chromosome segregation (Black et al., 2007; Lacoste et al., 2014). Others participate in specific transient processes, like the fundamental role of H2A.X in DNA damage response (van Attikum and Gasser, 2009). In a gene regulatory context, some histone variants like H2A.Z are associated with increased chromatin dynamics at actively transcribed genes while others tend to correlate with repressed states like the macroH2A histones (Buschbeck and Hake, 2017).

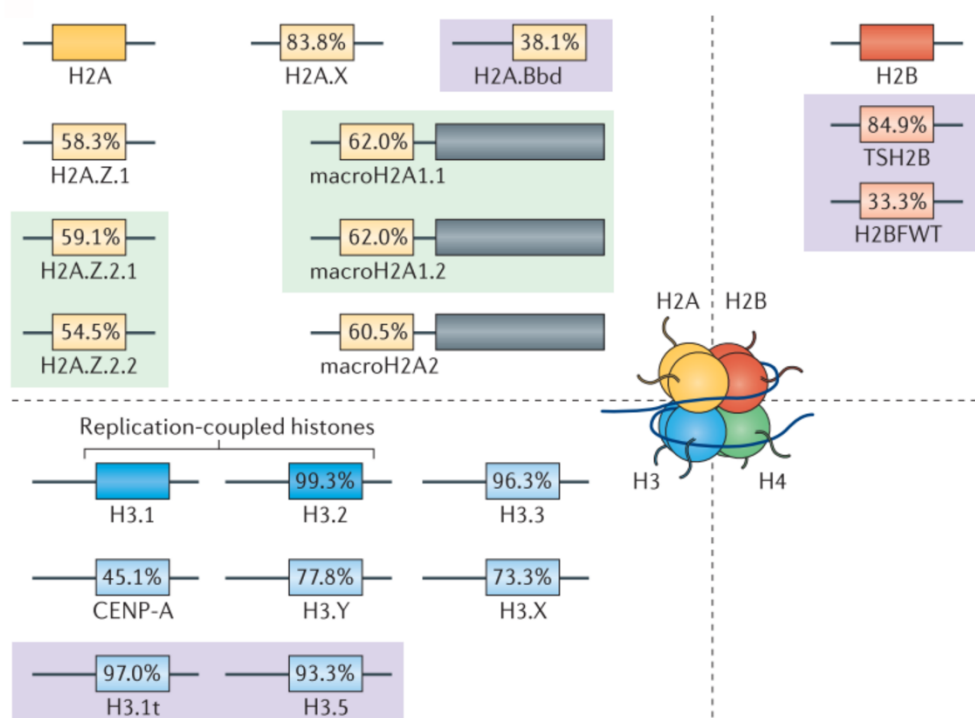


Figure 5. Histone variants in humans. Depiction of the human variants of histone H2A (yellow), H2B (red) and H3 (blue). Rectangles represent histone folds, while lines represent flexible histone tails. Percentages indicate the relative sequence identity of each variant relative to their replication-coupled counterpart (out of the two replication-coupled H3 histones, H3.1 was used to calculate this %). Testis-specific variants are highlighted in purple boxes, while alternative splicing isoforms by green boxes. No variants of H4 (green) have been identified in humans. Figure taken from Buschbeck and Hake, 2017.

Like replication-coupled histones, histone variants can also be subject to PTMs, expanding the functional diversity of this protein family (**Figure 6**) (Corujo and Buschbeck, 2018). Some PTMs are shared by histone variants and their replication-coupled counterpart while others are specific for a certain histone type. This is mainly due to sequence dissimilarities between replication-coupled histones and histone variants and differential recognition by the enzymatic complexes that catalyze the deposition of PTMs.

Introduction

While the function of some variant-specific PTMs remains poorly characterized, others like the phosphorylation and ubiquitination of H2A.X in the organization of the DNA damage response have been extensively studied (van Attikum and Gasser, 2009).

The expression of histone variants is dynamically regulated during embryonic development and they function in cell fate transitions and reprogramming (Buschbeck and Hake, 2017; Hoghoughi et al., 2017). In cancer, histone variants and their regulatory factors are found altered both transcriptionally and by mutations which can affect DNA damage repair, genomic stability and the transcription of oncogenes and tumor suppressor genes (Buschbeck and Hake, 2017; Corujo and Buschbeck, 2018).

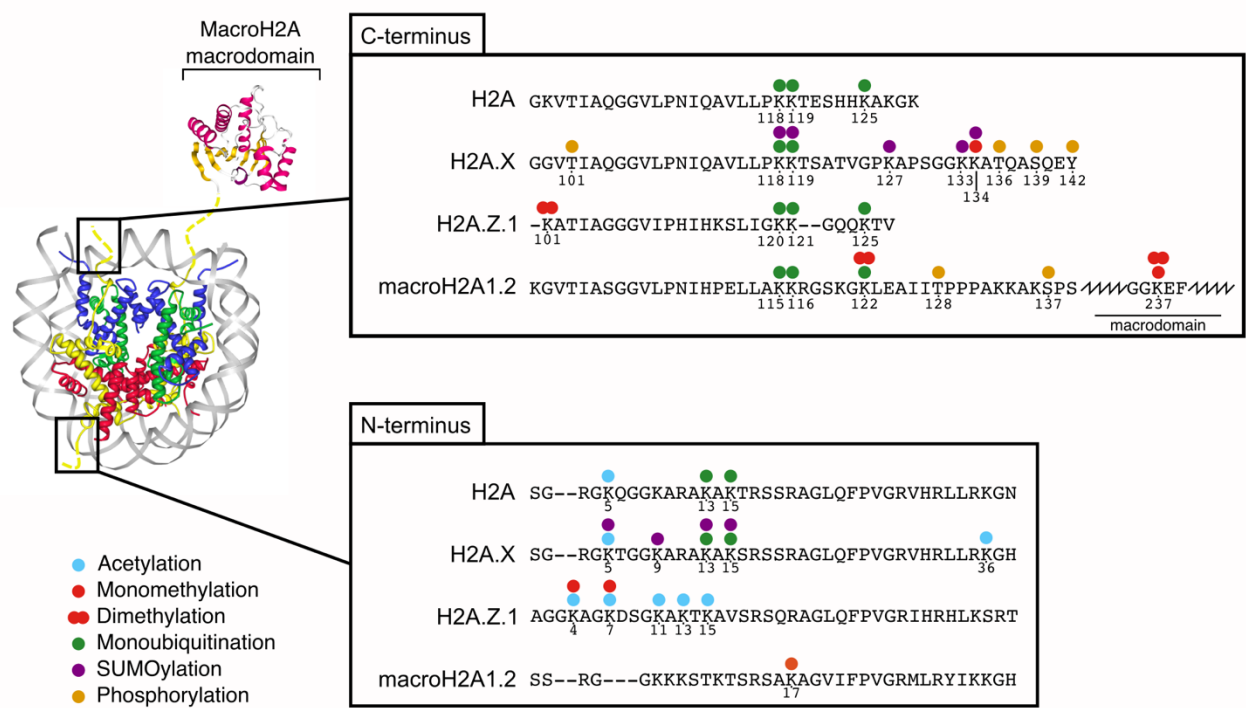


Figure 6. Amino acid sequence of the N- and C-termini of human H2A variants and their post-translational modifications. Alignment of human H2A type 1 (NCBI ID accession number NP_003501.1), H2A.X (NP_002906.1), H2A.Z.1 (NP_002907.1) and macroH2A1.2 (NP_004884.1). The crystal structure of a nucleosome containing a macroH2A histone is depicted. Histone H2A is colored in yellow, H2B in red, H3 in blue, H4 in green and DNA in grey. Dashed lines represent the H2A tails. The macromodomain of macroH2A is colored by secondary structure (α -helices in magenta and β -sheets in orange). The picture of the nucleosome is based on protein data bank ID 3REH (Wu, 2011) and generated with ProteinWorkshop (Moreland et al., 2005), the picture of the macroH2A macromodomain is based on PDB ID 2FXK (Kustatscher et al., 2005) and generated with NGL viewer (Rose and Hildebrand, 2015). Figure from Corujo and Buschbeck, 2018.

2. MacroH2A histone variants

2.1. The macroH2A family of histone variants

MacroH2A histones are the most divergent histone variants in terms of sequence and structure when compared to its replication-coupled counterpart. MacroH2A proteins are widespread and highly conserved in vertebrates, but recent work has evidenced their presence in an invertebrate organism and in the holozoan *Capsaspora owczarzaki*, suggesting a more ancestral evolutionary origin of macroH2A than initially thought (Rivera-Casas et al., 2016). There are three different macroH2A proteins in vertebrates: macroH2A1.1 and macroH2A1.2 arise by alternative splicing of a single gene (*H2AFY*) while macroH2A2 is encoded by a separate gene (*H2AFY2*) (**Figure 7A**) (Chadwick and Willard, 2001; Costanzi and Pehrson, 2001; Pehrson and Fried, 1992). In my thesis, I will use macroH2A to refer to all vertebrate macroH2A proteins and macroH2A1 to refer collectively to the splice isoforms macroH2A1.1 and macroH2A1.2 when it is not possible or necessary to distinguish them. Otherwise, macroH2A1.1, macroH2A1.2 or macroH2A2 will be used to refer to a specific protein.

MacroH2A proteins have a characteristic tripartite structure: in addition to an N-terminal histone fold, they contain a linker region and a large macro domain in the C-terminus (**Figure 7A**) (Chakravarthy et al., 2005a; Kustatscher et al., 2005). The histone fold of macroH2A shares around 65% sequence identity and the main structural features with replication-coupled H2A (Changolkar and Pehrson, 2002). The linker domain itself is a highly basic unstructured region that protrudes outside of the compact histone core placing the macro domain in a very accessible position (**Figure 7B**) (Chakravarthy et al., 2005a). Considering that the linker region and the macro domain together account for two times the molecular weight of the histone fold, the presence of macroH2A in a nucleosome can be seen as a drastic modification of its features.

The major peculiarity of macroH2A histones is the presence of its C-terminal macrodomain. It is a globular domain of about 25kDa in size composed of a seven-strand β -sheet sandwiched by five α -helices and presents an hydrophilic pocket (Chakravarthy et al., 2005a; Kustatscher et al., 2005). Macrodomains are highly conserved domains that, besides macroH2A, are present in non-chromatin components and can bind NAD⁺ derived metabolites such as ADP-ribose and O-acetyl-ADP-ribose (Posavec et al., 2013). All three macroH2A histones share the same overall domain structure, but differences in

Introduction

the sequence of the macrodomain provoke changes in the size and hydrophobicity of the macrodomain pocket (**Figure 7C**). Due to this differences, only the macrodomain of macroH2A1.1 is able to bind ADP-ribose and ADP-ribosylated proteins (Kustatscher et al., 2005). This particularity allows macroH2A1.1 to bind and, in some contexts, inhibit the activity of the enzyme poly-ADP-ribose polymerase 1 (PARP1), thereby regulating PARP1-dependent processes such as DNA-damage repair (Timinszky et al., 2009), stress response (Ouararhni et al., 2006) and transcriptional regulation (Chen et al., 2014). Notably, the splicing event between the isoforms macroH2A1.1 and macroH2A1.2 involves a mutually exclusive exon usage that determines the structure of the macrodomain pocket and subsequent binding of ADP-ribose (Kustatscher et al., 2005). While some macrodomains have ADP-ribosyl-hydrolase enzymatic activity, it is not the case for macroH2A1.1 (Jankevicius et al., 2013).

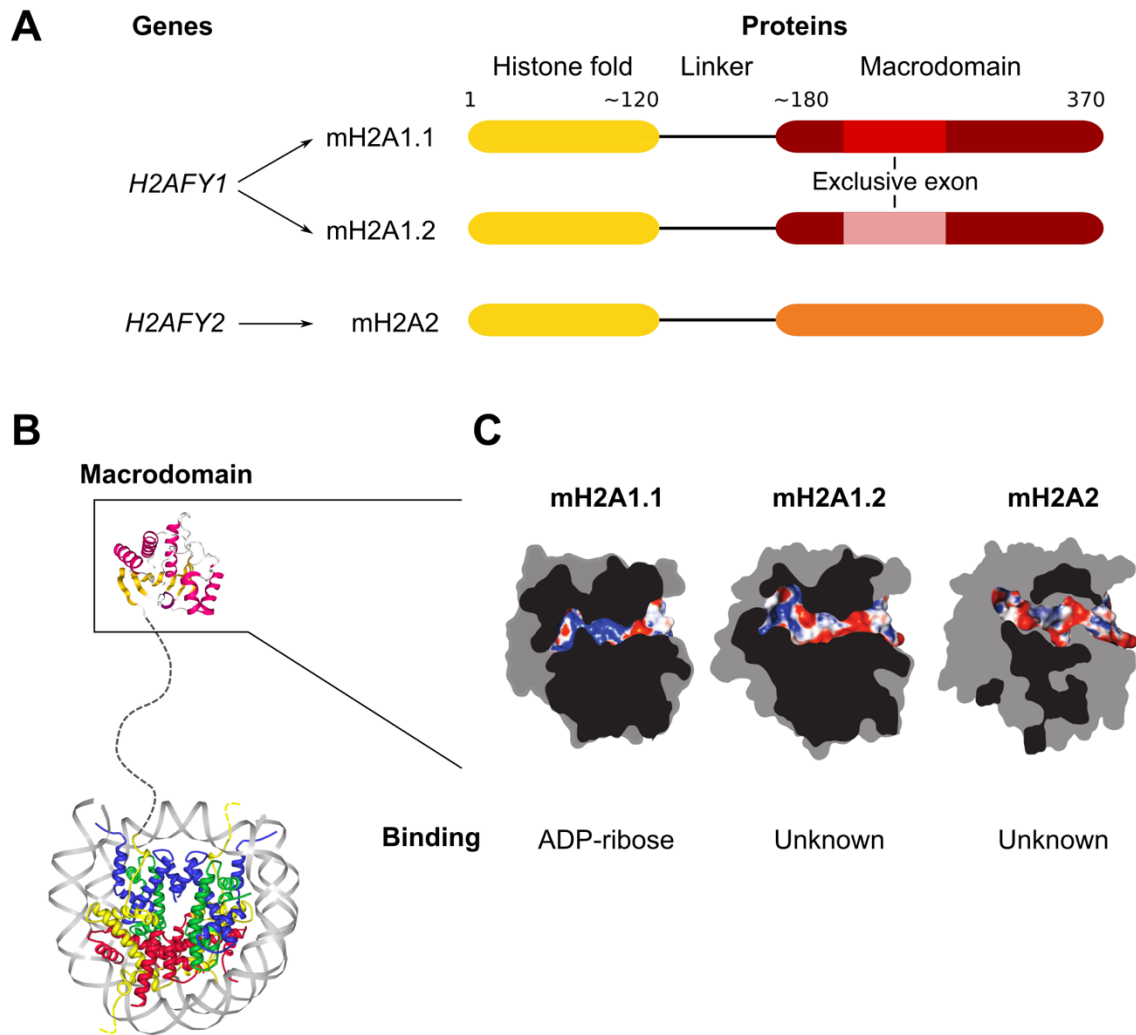


Figure 7. The family of macroH2A histone variants. **(A)** Schematic representation of the three macroH2A human histone variants encoded by the genes *H2AFY* and *H2AFY2*. The numbers refer to the amino acid number in the linear sequence. **(B)** Representation of the structure of a nucleosome containing a macroH2A histone variant. H2A and the histone fold of macroH2A are colored in yellow, H2B in red, H3 in blue, H4 in green and DNA in grey. The linker region of macroH2A is depicted as a dashed blacked line as no crystal structure has been achieved for this part of the protein. The macrodomain of macroH2A is colored by secondary structure (α -helices in magenta and β -sheets in orange). The picture of the nucleosome is based on protein data bank ID 3REH (Wu et al., 2011) and generated with ProteinWorkshop (Moreland et al., 2005), the picture of the macroH2A macrodomain is based on PDB ID 2FXK (Kustatscher, 2005) and generated with NGL viewer (Rose and Hildebrand, 2015). **(C)** Cross-section of the surface representation of the macrodomains of macroH2A proteins showing the structure and properties of their pockets. The interior of the pocket is colored to the surface charges: red for acidic, blue for basic and white for non-charged. The protein interior is colored in black, the rest of the protein is shaded in grey. This panel is adapted from Posavec, 2013.

2.2. Molecular and physiological functions of macroH2A

Multiple studies have implicated macroH2A proteins in development, cellular differentiation, somatic cell reprogramming and cancer (**Figure 8**). In general, macroH2A histones are considered to be stabilizing factors that correlate with differentiated states, but the precise molecular mechanism by which they act on chromatin and their role in chromatin regulation is still a complex and incomplete picture. In particular, many roles of macroH2A seem to be context-dependent and present functional variability between the three macroH2A proteins.

2.2.1. Nucleosomal stability and structure

In vitro, macroH2A preferentially forms “hybrid nucleosomes” containing one replication-coupled H2A and one macroH2A histone (Chakravarthy and Luger, 2006). The alterations introduced to the structure of the nucleosome core by the incorporation of macroH2A are based on a four amino acid stretch limited to the L1-L1 interface that is the only point of contact between the two H2A-H2B dimers (Chakravarthy and Luger, 2006; Chakravarthy et al., 2005a). *In vitro*, histone octamers containing macroH2A have increased stability in terms of ionic strength requirement and are resistant to the H2A-H2B dimer exchange by the histone chaperone NAP-1 (Chakravarthy and Luger, 2006). Studies based on molecular dynamic simulations have reached similar conclusions regarding increased nucleosome stability of hybrid macroH2A nucleosome particles (Bowerman et al., 2019). The linker region protrudes from the nucleosome core in very close proximity to the dyad axis (Chakravarthy et al., 2005a) and the presence of macroH2A could interfere with the binding of a transcription factor when its binding site was close to the dyad axis (Angelov et al., 2003).

2.2.2. Genomic localization and transcriptional regulation

Initial discovery of macroH2A identified its association with the inactive X chromosome (Costanzi and Pehrson, 1998; Mermoud et al., 1999), although it is not required for X chromosome inactivation (Tanasijevic and Rasmussen, 2011). Later studies have shown its broad distribution in mammalian autosomal chromatin, mostly in the form of large domains that can span hundreds of kilobases (Changolkar et al., 2010; Gamble et al., 2010).

MacroH2A-enriched regions generally correlate with the presence of H3K27me3 (Buschbeck et al., 2009; Gamble et al., 2010; Gaspar-Maia et al., 2013), a hallmark of

polycomb-mediated gene silencing (Schuettengruber et al., 2017). *In vitro* studies supported a role of macroH2A in transcriptional repression by showing that macroH2A interfered with SWI/SNF nucleosome remodeling and transcriptional initiation (Angelov et al., 2003; Chang et al., 2008; Doyen et al., 2006). Despite being depleted in active genes and being generally associated to transcriptional repression, macroH2A is in some cases required for signal-induced gene expression (Barrero et al., 2013a; Creppe et al., 2012; Gamble and Kraus, 2010; Ouarrhni et al., 2006). Moreover, macroH2A has been proposed to maintain robust gene expression patterns by reducing transcriptional noise (Lavigne et al., 2015). Interestingly, a fraction of macroH2A1.1 is present at H2B acetylated chromatin regions where it interacts with PARP1 and can either promote or repress gene expression (Chen et al., 2014; Ruiz and Gamble, 2018).

The molecular machinery responsible for macroH2A incorporation into chromatin is unknown, but a recent study shows that macroH2A2 deposition happens genome-wide and is later evicted from transcribed regions in a “pruning” manner that requires the FACT complex (Sun et al., 2018). Another study followed the incorporation of macroH2A into chromatin after DNA replication and found it targets the same large domains where it was previously enriched (Sato et al., 2019). Additionally, the histone fold of macroH2A is sufficient to target its deposition to H3K27me₃-marked chromatin (Chadwick et al., 2001; Nusinow et al., 2007a; Ruiz and Gamble, 2018).

2.2.3. DNA repair of double-strand breaks

MacroH2A histones have been implicated in the chromatin rearrangement occurring in response to DNA damage in the cell. MacroH2A1.1 is quickly recruited to DNA damage sites upon PARP1 activation (Timinszky et al., 2009), while macroH2A1.2 accumulation occurs later and promotes double-strand break repair by homologous recombination (Khurana et al., 2014). Genomic regions particularly sensitive to replication stress are termed fragile sites and are a frequent source of genomic instability in cancer cells (Barlow et al., 2013; Glover et al., 2017). A recent study shows that macroH2A1.2 accumulates at fragile genomic regions and promotes BRCA1 recruitment at stalled replication forks, protecting against replication stress and damage-induced senescence (Kim et al., 2018a).

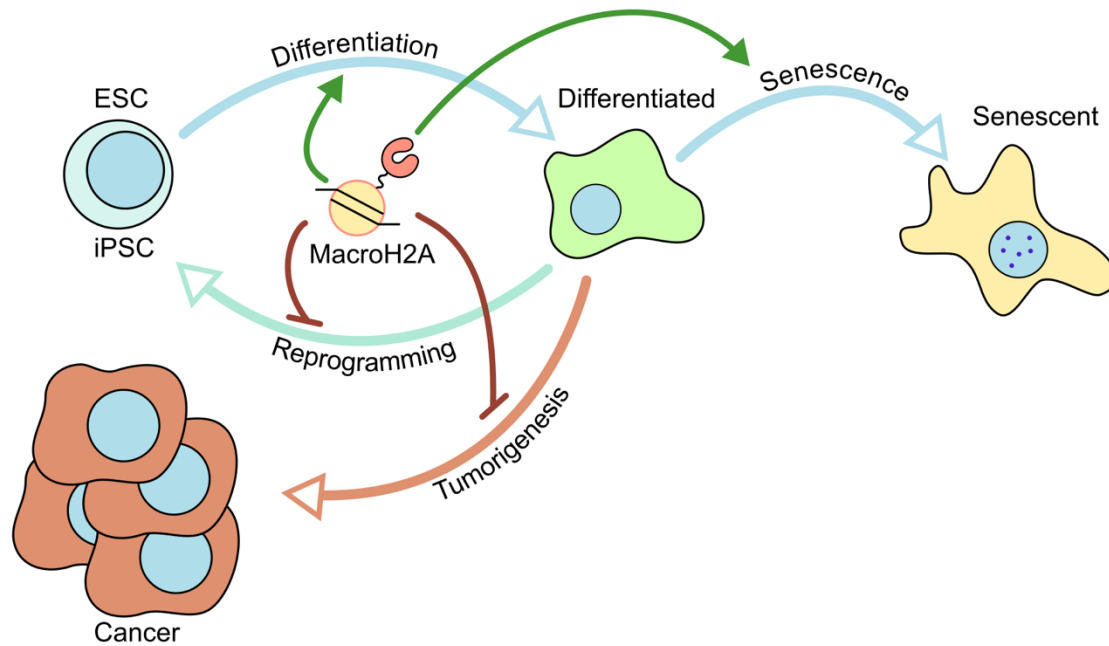


Figure 8. MacroH2A roles in cellular physiology. Schematic summary of macroH2A functions in different cellular processes. MacroH2A histones are generally correlated with the stabilization of differentiated states, act as barriers toward reprogramming and are considered to be mostly tumor-suppressors. MacroH2A1.1 in particular is implicated in paracrine senescence. Note that relevant differences between different macroH2A proteins are omitted for simplification.

2.2.4. Positive factor in differentiation and senescence

Mice lacking macroH2A are viable but have a retarded growth and metabolic alterations (Pehrson et al., 2014), while macroH2A-deficient zebrafish embryos display several developmental defects (Buschbeck et al., 2009). During mouse early embryonic development, macroH2A is first detected at the 8-cell stage (Nashun et al., 2010). Moreover, macroH2A levels generally increase during pluripotent stem cell differentiation and promote the differentiation of embryonic stem cells (Barrero et al., 2013a; Creppe et al., 2012). Interestingly, macroH2A1.2 is the predominant splicing isoform expressed in undifferentiated cells and cancer cell lines, while macroH2A1.1 expression correlates with differentiation (Posavec-Marjanović et al., 2017; Sporn and Jung, 2012). In the converse process, macroH2A histones function as an epigenetic barrier to induced pluripotency by somatic cell nuclear transfer and Yamanaka factors (Barrero et al., 2013b; Gaspar-Maia et al., 2013; Pasque et al., 2011b, 2011a, 2012).

Cellular senescence is a stable state of cell cycle arrest that happens in response to various stresses and involves a complex phenotype that includes morphological, metabolic and transcriptional programs (Salama et al., 2014). In senescent cells, macroH2A is enriched in senescence-associated heterochromatic foci and is frequently used as a marker to identify these chromatin structures (Zhang et al., 2005a). In particular,

macroH2A1.1 promotes the senescence-associated secretory phenotype in paracrine senescence triggered by oncogenes (Chen et al., 2015).

2.2.5. Context-dependent tumor suppression

MacroH2A histones are generally regarded as having a tumor suppressive function in cancer. MacroH2A expression decreases as the disease progresses in melanoma, bladder cancer and anal neoplasms (Hu et al., 2016; Kapoor et al., 2010; Kim et al., 2013). Knocking-down macroH2A results in more aggressive teratomas, melanoma, breast cancer and bladder cancer (Creppe et al., 2012; Dardenne et al., 2012; Kapoor et al., 2010; Park et al., 2016) while its overexpression reduces proliferation in several cancer cell lines (Dardenne et al., 2012; Kapoor et al., 2010; Novikov et al., 2011). Furthermore, down-regulation of macroH2A results in increased metastatic potential of melanoma (Kapoor et al., 2010).

However, there are important differences in the effect of the different isoforms of macroH2A in cancer (Cantariño et al., 2013). In particular, the expression level of macroH2A1.1 is inversely correlated with proliferation and its low levels are a marker for poor prognosis and metastatic potential in lung and colon cancer (Li et al., 2016; Novikov et al., 2011; Sporn and Jung, 2012; Sporn et al., 2009). In triple negative breast cancer, however, macroH2A1.1 expression correlates with poor survival and EMT (Lavigne et al., 2014). In contrast, macroH2A1.2 expression is generally higher in highly proliferative cancer cell lines (Chen et al., 2014) and increases migration, invasion and growth in breast cancer (Dardenne et al., 2012; Li et al., 2012). Nonetheless, there is evidence that supports a tumor suppressive role of macroH2A1.2 in melanoma and bladder cancer (Kapoor et al., 2010; Park et al., 2016) which points out that macroH2A1.2 function might be largely context-dependent. The splicing factors MBNL1 and QKI favor the expression of the macroH2A1.1 splice variant over macroH2A1.2 (Novikov et al., 2011; Sporn and Jung, 2012), while the RNA helicases Ddx5 and Ddx17 do the opposite (Dardenne et al., 2012). Expression of MBNL1 and QKI is altered in various cancer types (Danan-Gotthold et al., 2015). In particular, expression of QKI limits proliferation in gastric cancer at least in part through promoting macroH2A1.1 expression (Li et al., 2016) and lower macroH2A1.1 and QKI expression is associated with less differentiated tumors in prostate cancer (Vieira-Silva et al., 2019).

The function of macroH2A in cancer has been suggested to be related to the transcriptional regulation of oncogenes, tumor suppressors and cell cycle regulators.

Introduction

MacroH2A represses the expression of CDK8 in melanoma cells which limits their proliferative potential (Kapoor et al., 2010; Lei et al., 2014). MacroH2A1 represses rDNA gene clusters that are upregulated in highly proliferative cells (Cong et al., 2014). MacroH2A1.2 specifically upregulates the ERBB2 gene which functions as a potent oncogene in breast and other cancers (Li et al., 2012). Interestingly, a higher macroH2A1.1 to macroH2A1.2 ratio positively regulates the expression of the tumor suppressor SOD3 (Dardenne et al., 2012). Inhibition of senescence in colorectal cancer cells by the transcription factor ZEB1 requires the repression of macroH2A1 expression (De Barrios et al., 2017). In a colorectal cancer cell line, however, macroH2A1 together with DNA methylation lead to the repression of the cell cycle inhibitor p16 which results in higher proliferation (Barzily-Rokni et al., 2011).

MacroH2A1 knock-down favors the acquisition of stem-like features in bladder cancer and hepatocellular carcinoma through the up-regulation of LIN28A transcription and NF- κ B signaling, respectively (Park et al., 2016; Lo Re et al., 2017). This observation is reminiscent of the function of macroH2A in limiting pluripotency and hindering somatic cell reprogramming and highlights its role as an epigenetic barrier to de-differentiation events.

In conclusion, macroH2A is in general a differentiation-promoting factor that limits the acquisition of malignant characteristics by cancer cells. While current evidence clearly supports a tumor suppressive role for macroH2A1.1 and macroH2A2, the function of macroH2A1.2 seems to depend greatly on the context of the particular cancer studied.

2.2.6. Impact on NAD⁺ metabolism

MacroH2A1.1 is necessary for proper mitochondrial function in skeletal muscle cells, maintaining cellular NAD⁺ homeostasis through the inhibition of PARP1 (Posavec-Marjanović et al., 2017). This observation highlights a global role of macroH2A1.1 independent of gene regulation. Some contradictory results exist regarding the metabolic consequences of knocking-out macroH2A in mice. One study showed that macroH2A1 knock-out (KO) mice displayed a pre-diabetic phenotype (Changolkar et al., 2007). However, another study reports the opposite effect where macroH2A1 KO mice on high-fat diet showed increased leanness and glucose tolerance (Sheedfar et al., 2015). Studies based on overexpression showed that exogenous expression of macroH2A1.1 reduced lipid accumulation in liver cancer cell lines (Pazienza et al., 2014), while macroH2A1.2 expression in mice inhibited adipogenesis (Pazienza et al., 2016).

3. Heterochromatin and the three dimensional organization of the genome

3.1. Heterochromatin

The term “heterochromatin” was coined by Emil Heitz in 1928 to refer to his observation of strongly stained regions of the nucleus in interphase cells under the microscope when using chromatin stains (Passarge, 1979). In contrast, he referred to lightly stained regions as “euchromatin” and, despite not knowing of the molecular composition of chromatin at the time, he hypothesized that “euchromatin is genically active, heterochromatin genically passive”. The later development of electron microscopy (EM) provided a similar distinction based on the differential electron density in various parts of the nucleus. In EM images, heterochromatin is visualized as dark regions mostly found around the nucleolus and close to the periphery of the nucleus, revealing a clear spatial compartmentalization (**Figure 9**). Although Heitz studies were based on a cytological definition, the terms “heterochromatin” and “euchromatin” are still commonly used nowadays at a molecular level to broadly divide eukaryotic chromatin into two major compartments.

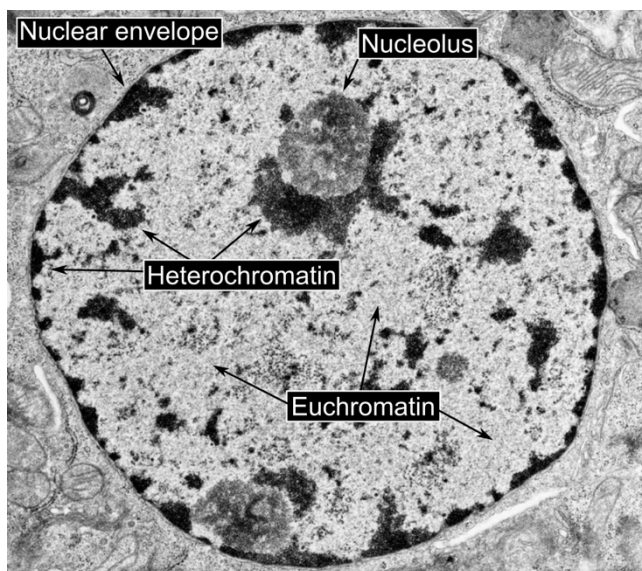


Figure 9. Transmission electron microscopy image of an eukaryotic cell nucleus. Darkness correlates with electron density. Regions of heterochromatin and euchromatin, as well as a nucleolus and the nuclear envelope, are indicated. Image adapted from the Yale Medical Studies Histology Image Gallery (Peter Takizawa).

Heterochromatin is mostly transcriptionally repressed and exhibits a high degree of compaction, in contrast to the more “open” euchromatin that tends to contain actively transcribed genes. Several decades of study have revealed major molecular differences between euchromatin and heterochromatin involving chromatin folding, DNA methylation, histone PTMs, and associated protein complexes. Additionally, heterochromatin is characterized by a late replication timing in S-phase (Gilbert, 2002). Two major subtypes of heterochromatin are usually considered: constitutive heterochromatin

Introduction

and facultative heterochromatin (**Figure 10**). Constitutive heterochromatin forms the major heterochromatic blocks that are transcriptionally repressed in all cell types of an organism, while facultative heterochromatin contains silenced genes that can be activated in response to signals and varies between cell types and during development.

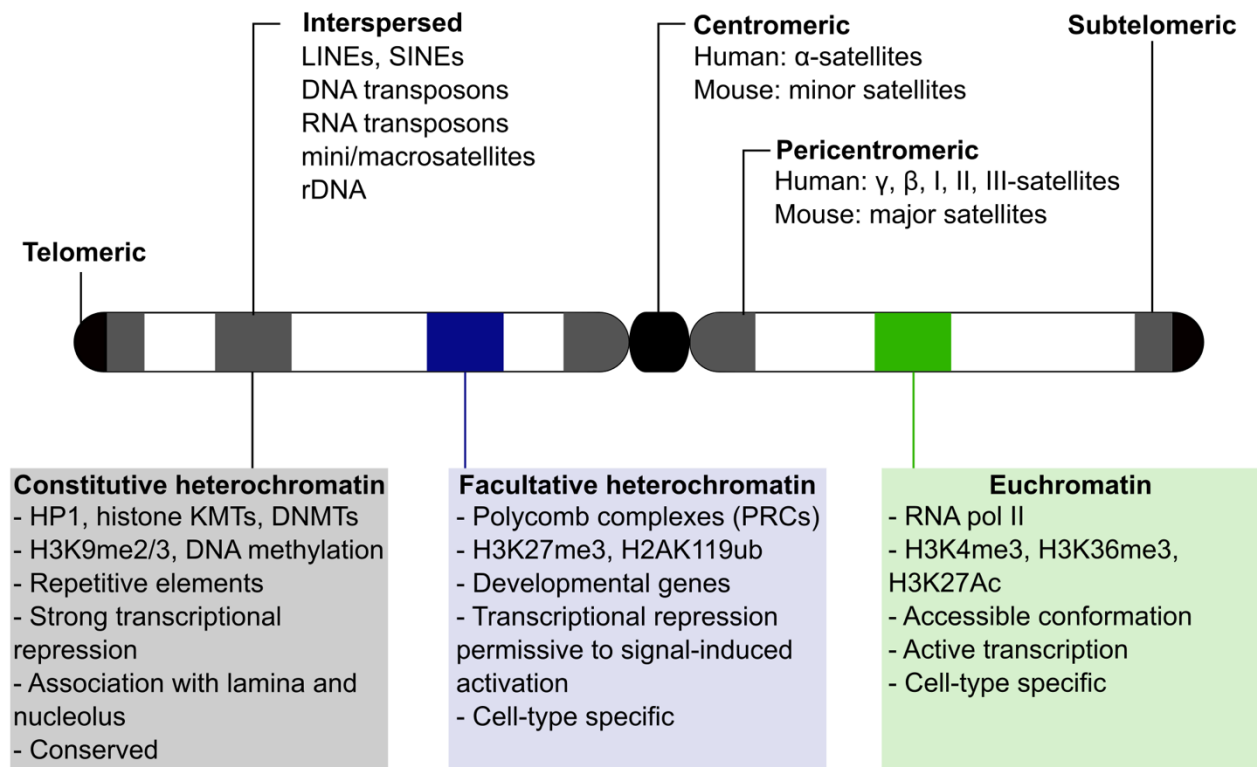


Figure 10. Repetitive elements in the mammalian genome and major features of heterochromatin and euchromatin. On the top, Schematic representation of the major classes of repetitive elements found in mammalian genomes and their location in the chromosome (inspired by Saksoouk et al., 2015). On the bottom, major features associated with genomic regions forming constitutive heterochromatin, facultative heterochromatin or euchromatin. Only the centromeric and telomeric chromosomal positions are meaningful, the blocks in the chromosomal arms are positioned arbitrarily for illustrative purposes.

3.1.1. Constitutive heterochromatin and repetitive elements

Constitutive heterochromatin is a highly conserved feature of eukaryotic chromatin from yeast to vertebrates, particularly in centromeric and telomeric chromosomal domains. Indeed, many insights into the conserved mechanisms of heterochromatin formation and function have been obtained in model organisms (Allshire and Madhani, 2017). It encompasses gene-poor genomic regions and is considered to have little variability between cell types in multicellular organisms. Constitutive heterochromatin contains most of the repetitive elements of the genome, which are estimated to represent between 50 and 70% of the human genomic sequence (Padeken et al., 2015). These repetitive elements include “structural” telomeric and centromeric domains, tandem satellite repeats found in subtelomeric and pericentromeric regions, and repeats

interspersed in euchromatic regions mostly composed of DNA transposons and retrotransposons (**Figure 10**) (Biscotti et al., 2015). Despite their abundance in the genome and due to their repetitive nature, these sequences remain largely excluded of the current genome assemblies and pose major challenges for experimental approaches based on high-throughput sequencing (Miga, 2015). A particularly distinct group of repeats is composed of the ribosomal DNA (rDNA) genes, which encode for the ribosomal RNAs that act as essential elements of the ribosomes. The rDNA repeats are highly transcriptionally active, in fact ribosomal RNAs are the most abundant RNA in proliferating cells, but a fraction of rDNA genes is kept silent in the form of constitutive heterochromatin (McStay and Grummt, 2008).

Many enzymatic activities and chromatin associated proteins are involved in the formation and function of constitutive heterochromatin (**Figure 11**). Methylation of H3K9, and in particular H3K9me3, is the most conserved hallmark of constitutive heterochromatin and is upstream of many heterochromatin characteristics. In mammalian cells, a series of SET-domain containing methyltransferases catalyze the deposition of this mark: G9a and GLP are responsible for mono- and dimethylation of H3K9 (Tachibana et al., 2002, 2005), while SETDB1, SETDB2, SUV39H1 and SUV39H2 generate di- and trimethylated H3K9 (Falandry et al., 2010; O'Carroll et al., 2000; Rea et al., 2000; Schultz et al., 2002). HP1 proteins bind H3K9me3 through their chromodomain and are essential for heterochromatin establishment and maintenance (Bannister et al., 2001; Lachner et al., 2001). HP1 proteins can oligomerize through their chromoshadow domain (Canzio et al., 2011; Cowieson et al., 2000) and contain an unstructured hinge that can directly interact with nucleic acids (Meehan et al., 2003), providing a mechanism of heterochromatin compaction and compartmentalization that has been recently proposed to involve phase separation properties (Larson et al., 2017; Strom et al., 2017). Other chromatin regulators including nucleosome remodeling complexes and transcription factors participate in heterochromatin formation and maintenance (Becker and Hörz, 2002; Bulut-Karslioglu et al., 2012). Importantly, most studies on heterochromatin nucleation, spreading and maintenance have been performed in *S. pombe* yeast cells, but most of the participant proteins have orthologs in vertebrates and the basic principles are considered to be conserved.

Besides H3K9me3, constitutive heterochromatin is characterized by global histone deacetylation (Taddei et al., 2001) and is associated with other histone PTMs including H3K20me2/3 (Jorgensen et al., 2013). Notably, DNA methylation is also extensively found

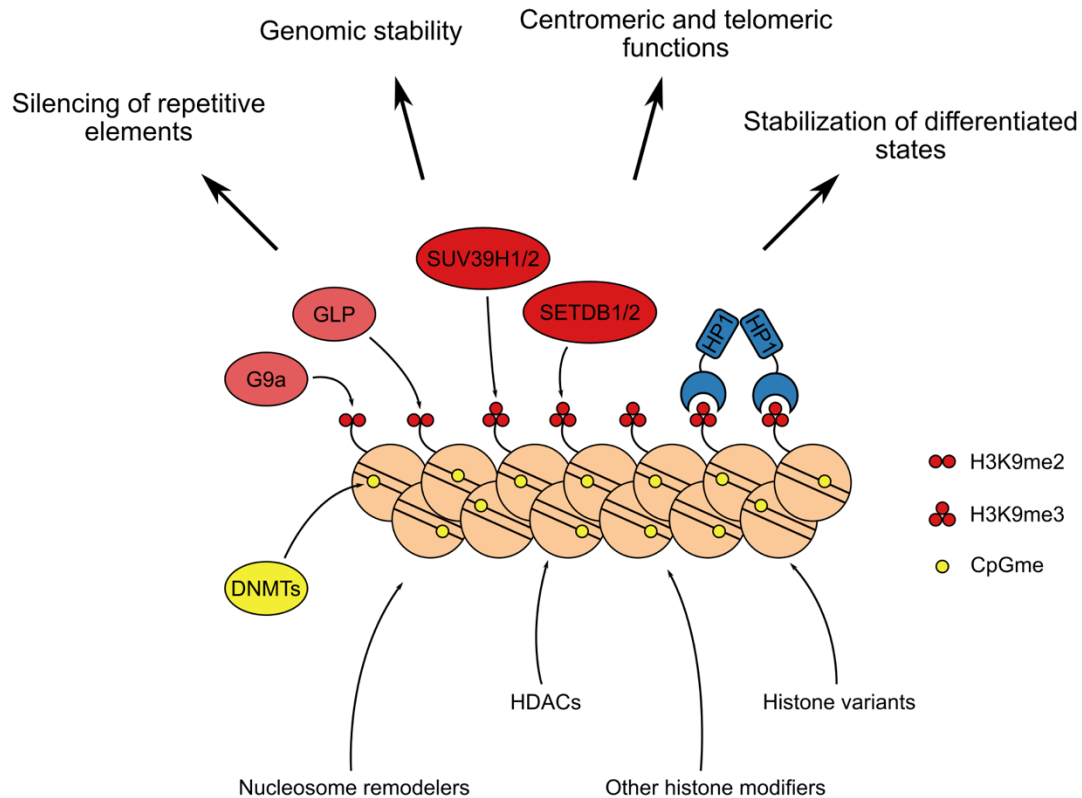


Figure 11. Constitutive heterochromatin formation and functions. Schematic representation of the major factors involved in the formation of constitutive heterochromatin and some of the main functions of this type of heterochromatin (in mammals, although the principles are mostly conserved in eukaryotes). A series of histone methyltransferases catalyse the deposition of H3K9me3 on chromatin, which serves as a binding site for HP1 proteins that can dimerize and oligomerize to promote compaction. Other factors such as DNA methyltransferases (DNMTs), histone deacetylases (HDACs), nucleosome remodeler complexes and histone participate in heterochromatin structure and function. Figure inspired in Saksouk, 2015.

and required at heterochromatic regions, as removal of DNA methylation reduces H3K9me3 levels and disrupts pericentromeric architecture (Saksouk et al., 2014). ATP-dependent chromatin-remodeling complexes are also involved in the compaction of constitutive heterochromatin and its higher-order structure. The NuRD, NoRC and HELLS complexes promote chromatin compaction and histone deacetylation in heterochromatic regions (Saksouk et al., 2015). Certain histone variants are also incorporated at repetitive constitutive heterochromatin regions. Most notably, the histone H3 variant CENP-A is essential for the structure of the centromere and the assembly and function of the kinetochore in chromosome segregation (Sharma et al., 2019), while H3.3 deposition by ATRX/DAXX contributes to the establishment of heterochromatin at telomeric, pericentromeric and retroviral elements (Voon and Wong, 2016).

Although they are a source of genetic variability, the activity of transposable elements poses a threat to the integrity of the genome as they can introduce deleterious mutations. Additionally, the highly repetitive nature of pericentromeric regions is propense

to aberrant recombination. Thus, preservation of genome stability through the transcriptional silencing of repetitive elements is a major role of constitutive heterochromatin (Janssen et al., 2018). In addition, proper heterochromatin structures are essential for centromeric and telomeric function and proper chromosome segregation (García-Cao et al., 2004; Peters et al., 2001). Increased genomic instability is a hallmark of cancer that is proposed to provide cancer cells with increased plasticity to acquire malignant properties and confer tumor heterogeneity (Hanahan and Weinberg, 2011). Changes in the levels and distribution of H3K9me3 and HP1 as well as the aberrant expression of interspersed and pericentromeric repetitive elements are associated with cancer progression (Dialynas et al., 2008; Slee et al., 2012; Ting et al., 2011).

Besides silencing of repeats, H3K9me3-marked heterochromatin has also been involved in gene silencing mechanisms that stabilize differentiated states (Becker et al., 2016). Large heterochromatic blocks are acquired during differentiation to silence pluripotency and lineage-inappropriate genes and act as a barrier to reprogramming (Soufi et al., 2012; Wen et al., 2009).

In summary, constitutive heterochromatin is compact and transcriptionally repressed and is characterized by the presence of H3K9me3. It is well conserved between cell-types, keeps repetitive elements silent and maintains genomic stability.

3.1.2. Facultative heterochromatin and polycomb

Facultative heterochromatin contains mostly coding sequences and is found in gene-rich regions where it mediates transcriptional silencing. In contrast to constitutive heterochromatin, facultative heterochromatin regions can be highly variable between cell types and during development. It provides a dynamic mechanism of gene regulation by silencing genes but keeping them responsive to transcriptional activation by developmental signals. Facultative heterochromatin is a key element in cellular identity specification and maintenance of both pluripotent and lineage-specific gene expression programs.

The hallmark of facultative heterochromatin is the presence of polycomb repressive complexes and the histone PTMs H3K27me3 and H2AK119ub. In fact, genomic regions of H3K27me3 enrichment are usually termed “polycomb domains”, where H3K27me3 covers the promoter and body of silenced genes. PRCs are composed of the polycomb group of proteins (PcG) which mainly associate in two major complexes: PRC1 and PRC2. Both complexes have enzymatic activity and modify histone proteins: PRC1

Introduction

monoubiquitinates H2AK119 while PRC2 trimethylates H3K27. Initially identified in *Drosophila*, these complexes are functionally conserved in metazoans. Unless stated otherwise, in this section I will refer to the mammalian subunits and mechanisms of PRCs. PRCs are highly modular protein complexes with core components that include their catalytic activity that can be modulated by the presence of other subunits and associated proteins (**Figure 12**) (Di Croce and Helin, 2013).

The core PRC2 is formed by the association of the proteins EED, SUZ12 and EZH2 or EZH1 (Schuettengruber et al., 2017). EZH2 and EZH1 are the mutually exclusive enzymatically active components of the complex that can catalyze the mono-, di- and trimethylation of H3K27 through their SET domain (Cao et al., 2002; Jiao and Liu, 2015; Margueron et al., 2008; Shen et al., 2008). EED and SUZ12 are further required for the recruitment and activity of the complex (Cao and Zhang, 2004).

PRC1 complexes have a more heterogeneous composition and are divided into canonical and non-canonical PRC1 complexes (cPRC1 and ncPRC1, respectively). All PRC1 complexes share a core composed by a RING1 protein (RING1A or RING1B) and one of six PCGF subunits (PCGF1-6). RING1 proteins are E3 ubiquitin ligases that can ubiquitinate H2A at lysine 119 (Cao et al., 2005; Wang et al., 2004). cPRC1 is characterized by the presence of a CBX subunit (CBX2, CBX4, CBX6-8). CBX proteins contain a chromodomain that can bind H3K27me3 and each CBX subunit confers distinct gene-regulatory functions to the PRC1 complex (Di Croce and Helin, 2013). On the other hand, ncPRC1 complexes contain RYBP or YAF2 proteins and are further specified by the presence of other factors (Schuettengruber et al., 2017).

The classical model of PRC function proposes a sequential mode of action in which PRC2 is recruited to target genes and deposits H3K27me3 that is bound by PRC1 through its CBX subunit leading to ubiquitination of H2AK119. However, multiple studies have challenged this dogma, evidenced by the observation that the levels and distribution of H2AK119ub are not affected in PRC2-deficient mouse cells (Tavares et al., 2012). Moreover, ncPRC1 can be recruited to chromatin in the absence of PRC2 and H2AK11ub has been shown to recruit PRC2 (Blackledge et al., 2014; Tavares et al., 2012). The recruitment of PcG proteins to specific chromatin sites and target genes is a topic of current investigation for which locus-specific and generic targeting models involving transcription factors, CpG islands and ncRNAs have been proposed (Blackledge et al., 2015).

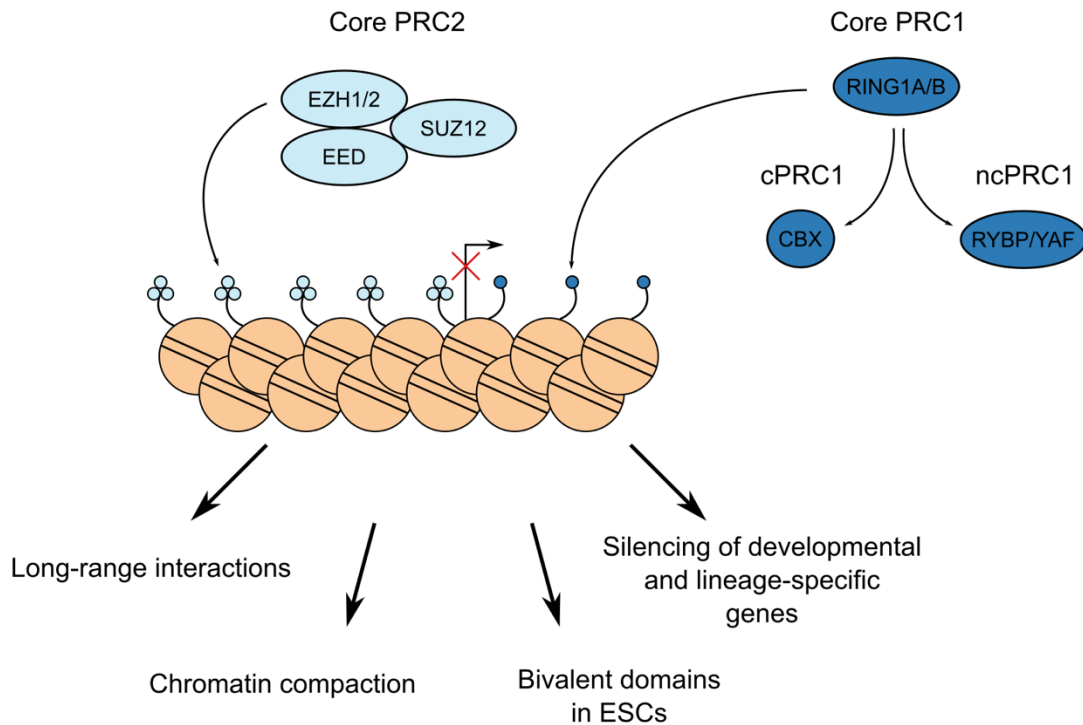


Figure 12. Polycomb repressive complexes and their function in chromatin. On the top, schematic and simplified composition of the PRC2, canonical PRC1 (cPRC1) and non-canonical PRC1 (ncPRC1). Both PRC2 and PRC1 can associate with many other proteins that further specify complex subtypes and functions that are not represented in this scheme. PRC2 and PRC1 bind and modify chromatin repressing gene expression, some of the major mechanisms and processes PRCs participate in are summarized on the bottom.

Different mechanisms for PRC-mediated gene repression have been proposed. H2AK119ub was thought to directly block transcription, but recent studies have demonstrated that this mark is dispensable for polycomb repression (Illingworth et al., 2015; Pengelly et al., 2015). Rather than a direct effect of histone PTMs on nucleosomal structure, one of the main PRC-mediated gene silencing mechanisms is thought to be the compaction of chromatin which reduces the accessibility of remodeler complexes and transcriptional machinery (Francis et al., 2004; Grau et al., 2011; Lau et al., 2017). Moreover, recent studies have highlighted the role of PRCs in spatial genome organization, where PcG proteins can mediate long-range contacts maintaining specific interaction networks of silenced genes (Entrevan et al., 2016).

PRCs are involved in the maintenance of pluripotency in embryonic stem cells (ESCs) by repressing many developmental regulators and lineage-specific genes (Boyer et al., 2006; Endoh et al., 2008; van der Stoop et al., 2008). Upon differentiation signals, the genome-wide distribution of PcG proteins mediates the acquisition of cell-type specific gene expression programs (Boyer et al., 2006). Remarkably, PRCs occupy “bivalent”

promoters of developmental regulators that contain both repressive H3K27me3 and activating H3K4me3 marks, keeping genes silent but highly responsive to activation signals (Bernstein et al., 2006; Harikumar and Meshorer, 2015b; Mas et al., 2018). PRCs also function in differentiation processes that occur in the adult organism. For instance, PRCs are necessary for proper hematopoiesis and alterations in PcG proteins are found in many hematological malignancies (Iwama, 2019; Park et al., 2003; Xie et al., 2014).

In conclusion, facultative heterochromatin is characterized by the presence of PRCs whose activity deposits H3K27me3 and H2AK119ub on chromatin. This type of heterochromatin is cell-type specific and maintains developmental and lineage-specific genes silent but sensitive to signal-induced activation.

3.2. Basic principles of 3D genome organization

Early microscopy observations already indicated the complexity of chromatin folding in the nucleus, with regions of varying density distinctly positioned in the nuclear space. Moreover, the vertebrate genome is characterized by long linear distances between target genes and *cis*-regulatory elements such as enhancers, hinting the necessity for particular conformations to mediate the effect. Over the past two decades, the study of the three dimensional (3D) organization of the genome has gained a lot of attention and is now recognized as a significant factor virtually in all chromatin functions including transcriptional regulation, DNA replication and repair.

The development of super-resolution fluorescence microscopy has overcome the diffraction limit and allowed to image higher-order chromatin structures in single fixed and live cells (Xu and Liu, 2019). In addition, chromosome conformation capture technologies have been a major breakthrough in the determination of chromatin interaction frequencies in the 3D-space (Sati and Cavalli, 2017). Chromosome conformation capture was initially developed by Dekker et al. (Dekker et al., 2002) and has served as the basis for many derivative assays, all based on the fixation of chromatin followed by enzymatic digestion of DNA and the generation of ligation fragments based on their three-dimensional proximity. “HiC” is a variation of chromosome conformation capture that uses high-throughput sequencing for the detection of such ligation events allowing the genome-wide determination of chromatin interactions.

These and other technical advancements have permitted the detailed study of the genome in 3D, revealing a complex multilayered “architecture” with hierarchical features

(Figure 13). These features are chromosome territories, A/B compartments, lamina-associated domains and nucleolus associated-domains, topologically associated domains and local interactions.

3.2.1. Chromosome territories

At the largest scale, individual chromosomes occupy distinct regions in the nuclear space which are termed “chromosome territories” **(Figure 13A)** (Cremer and Cremer, 2010). Initially described by the use of “chromosome painting” using fluorescent probes (Croft et al., 1999), chromosome territories showed a non-random radial distribution with gene-dense chromosomal segments positioned in the center of the nucleus and gene-poor heterochromatic regions located in the nuclear periphery (Boyle et al., 2001; Cremer et al., 2001). The mechanisms that form and maintain chromosome territories remain poorly understood, but attachment to the nuclear lamina seems to be an important factor (Bronshtein et al., 2016). Genome wide chromosome conformation capture (HiC) experiments also show preferential contacts between regions of the same chromosome (Lieberman-aiden et al., 2009). Although inter-chromosomal contacts happen at a much lower rate, a certain level of intermingling between chromosome territories exists and is proposed to have a function in gene regulation (Branco and Pombo, 2006; Szczepińska et al., 2019). These observations led to the organizational model where chromatin loops containing active genes colocalize around “transcription factories” enriched in RNA polymerase II and around nuclear speckles **(Figure 13B)** (Fraser and Bickmore, 2007). Recent studies have identified two distinct inter-chromosomal contact hubs around nuclear speckles and the nucleolus, respectively containing euchromatic and heterochromatic regions (Quinodoz et al., 2018). Still, the communication between functional regulatory elements on different chromosomes seems to be a rare event.

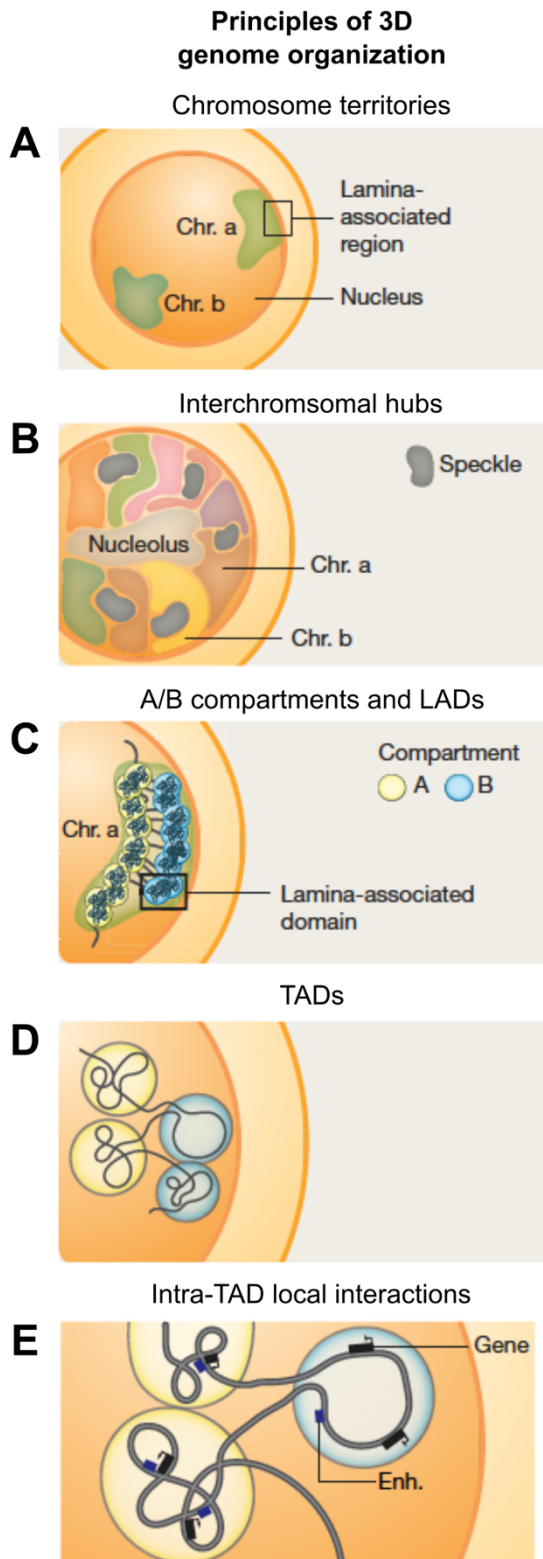


Figure 13. Basic principles of the 3D organization of chromatin. (A-E) Hierarchical chromatin organization structures ordered from the highest dimension (top) to the lowest (bottom). LADs, lamina-associated domains; TADs, topologically associating domains; Chr., chromosome; Enh., enhancer. Figure adapted from Stadhouders, 2019.

3.2.2. A/B compartments

Genome-wide chromosomal contacts as identified in HiC experiments broadly define the segregation of chromatin into the A and B compartments (Lieberman-aiden et al., 2009). Chromosomal segments in the A or B compartment preferentially interact with segments within the same compartment. The A compartment is considered to be the euchromatic fraction of the genome and is characterized by high gene density, transcriptional activity, chromatin accessibility and histone PTMs associated with gene activity. On the contrary, the B compartment is gene-poor, less accessible, transcriptionally inactive, is enriched in repressive histone marks and interacts with the nuclear lamina (**Figure 13C**). Microscopy studies mapping single chromosomes show that A/B compartments effectively separate in the nuclear space forming polarized structures (Wang et al., 2016). Higher resolution HiC datasets suggest that the A and B compartments can be further subdivided into six smaller compartments (A1, A2, B1, B2 and B3) (Rao et al., 2014). This subdivision is particularly interesting for the B compartment as it identified different types of heterochromatin: B1 is enriched in H3K27me3, indicating that it is formed by polycomb-repressed facultative heterochromatin; B2 is enriched in pericentromeric heterochromatin and interacts with the nuclear lamina and the nucleolus; and B3 is enriched at the nuclear

lamina but depleted at the nucleolus. Interestingly, the A/B compartmentalization defined in HiC recalls many of the features of chromosomal organization that were identified by the study of chromosome territories.

3.2.3. Domains associated with the nuclear lamina and the nucleolus

The nucleus is physically defined by its enclosure in the nuclear envelope, a double lipid membrane bilayer contiguous with the endoplasmic reticulum. Underneath the nuclear envelope lies a meshwork of type V intermediate filaments, the nuclear lamins, that form a structural scaffold called the nuclear lamina (de Leeuw et al., 2018). Electron microscopy imaging has long identified the accumulation of dense heterochromatin at the nuclear periphery (Fawcett, 1966). Recent advances in three-dimensional electron microscopy coupled with DNA labelling have also revealed a higher degree of chromatin packaging close to the nuclear lamina (Ou et al., 2017). Thus, the nuclear lamina is recognized as a repressive compartment enriched in heterochromatin.

Three genes, *LMNA*, *LMNB1* and *LMNB2* encode the four lamins forming the major part of this structure in mammals: lamin A and C, which are splice variants of the *LMNA* gene, lamin B1 and lamin B2 (de Leeuw et al., 2018). In addition to lamins, many additional associated proteins are found at the nuclear lamina, some of them being transmembrane proteins embedded in the inner membrane of the nuclear envelope, such as emerin, LAP2 β and the lamin B receptor (LBR) (**Figure 14A**) (Dobrzynska et al., 2016). While lamins are mostly located at the nuclear periphery, a fraction of lamin A/C and lamin B1 is found in the nucleoplasm interacting with euchromatic regions and participates in gene regulation (Cesarini et al., 2015; Dechat et al., 2010; Pascual-Reguant et al., 2018).

Genome-wide mapping of the association of lamins with chromatin has defined “lamin associated domains” (LADs). In mammalian cells, the genome is estimated to contain 1000-1500 LADs that can range from 10 kb up to 10 Mb with a median size of about 0.5 Mb, covering nearly 40% of the genome (Guelen et al., 2008; Peric-Hupkes et al., 2010). Interestingly, LADs generally overlap with the B compartment and recapitulate its characteristics: low gene density, enriched in repetitive elements and repressive histone PTMs, in particular H3K9me2/3 and H3K27me3 (Guelen et al., 2008; Harr et al., 2015; Kind et al., 2015). In a very similar way in which two major types of heterochromatin are defined, LADs have been subdivided in constitutive LADs (cLADs) and facultative or variable LADs (fLADs or vLADs) (**Figure 14B**). cLADs are consistent between different cell types, are AT-rich, accumulate LINE elements and are particularly gene-poor

Introduction

(Meuleman et al., 2013; Peric-Hupkes et al., 2010). Due to their high conservation, it is hypothesized that cLADs form a structural repressive backbone that contributes to the folding of chromosomes. On the other hand, vLADs are more gene-rich and contain lineage-specific genes that display changes in their lamina association and transcriptional activity during differentiation (Peric-Hupkes et al., 2010).

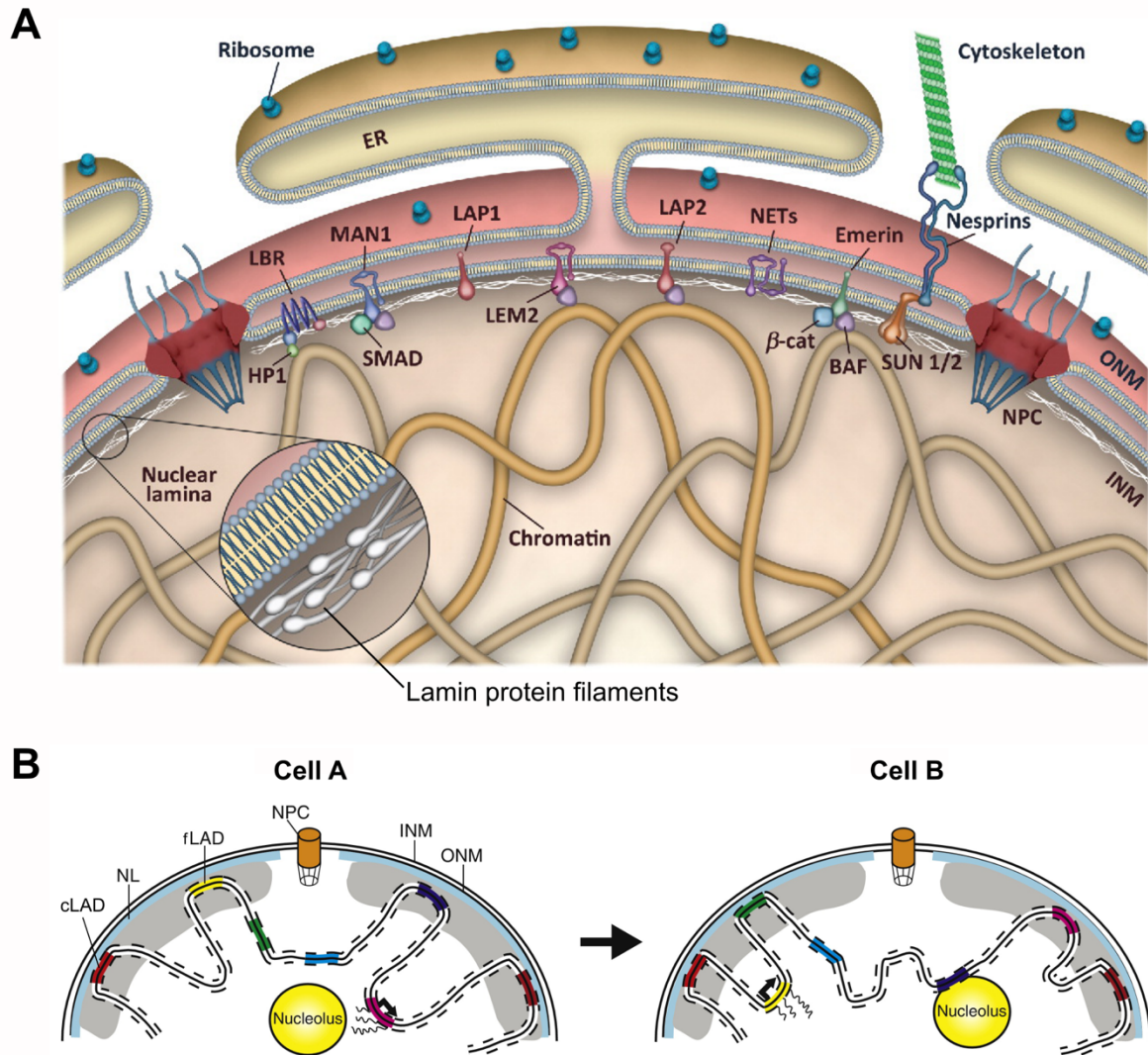


Figure 14. Chromatin interactions with the nuclear lamina. (A) Schematic representation of the nuclear envelope, the nuclear lamina and associated proteins. The nuclear envelope consists of an outer (ONM) and an inner nuclear membrane (INM), which are perforated by the nuclear pore complexes (NPC). The ONM is continuous with the endoplasmic reticulum (ER) and associated with ribosomes as well as nesprins and other proteins linking the nucleus to the cytoskeleton. The nuclear lamina is a network of lamin protein filaments underneath the INM that forms contacts with chromatin and INM proteins. The INM contains multiple nuclear envelope transmembrane (NET) proteins, of which only a few are depicted. Some of this proteins have been implicated in tethering chromatin to the nuclear lamina: LEM domain proteins (such as LAP2 β) interact with the chromatin-binding protein BAF and transcription factors, such as SMADs, β -catenin and others, whereas lamin B receptor (LBR) associates with heterochromatin protein 1 (HP1) and directly with modified histones. Panel adapted from Dobrzynska, 2016. **(B)** Schematic Model of dynamic chromatin-nuclear lamina interactions. cLADs show conservation between cells, while fLADs are more variable. In addition, heterochromatic regions can shuffle between the nuclear lamina (NL) and the nucleolus. In general, detachment from the repressive NL environment (in grey) permits transcriptional activity (arrows and wavy lines). Panel adapted from Amendola and van Steensel, 2014.

The specific interactions that mediate chromatin tethering to the nuclear lamina are not yet fully understood. Although lamins interact with chromatin *in vitro*, deletion of all lamins in mESC had little effect in the genome-wide contacts with the nuclear lamina (Amendola and Steensel, 2015). Deletion of H3K9 methyltransferases results in reduced LAD association with the nuclear lamina, which suggests a strong role for H3K9me2/3 (Bian et al., 2013; Harr et al., 2015; Towbin et al., 2012). Additionally, a series of transmembrane proteins embedded at the inner nuclear membrane (NETs, nuclear envelope transmembrane proteins), including emerin, LBR and LAP2 β have been implicated in anchoring chromatin to the nuclear lamina (**Figure 14A**) (Shevelyov and Ulianov, 2019; van Steensel and Belmont, 2017).

Although the nuclear lamina is considered to be a repressive environment, the causal relation between gene positioning and transcriptional activity is not fully understood nor consistent. The study of single cells revealed a stochastic re-positioning of LADs after mitosis where detachment from the lamina correlated with the elongation mark H3K36me3 (Kind et al., 2013a). Artificial tethering of genomic loci to the nuclear periphery caused reduced expression in some but not all tested genes (Finlan et al., 2008; Kumaran and Spector, 2008; Reddy et al., 2008). Although the majority of genes in LADs are not expressed, around 10% are transcriptionally active (Guelen et al., 2008). A recent study has identified intrinsic characteristics of promoter sequences that can lead a gene to “escape” the repressive environment in LADs (Leemans et al., 2019). Together, the current understanding is that the nuclear periphery effectively forms a repressive environment but that stochastic interactions, local DNA features and chromatin states inside LADs can still overcome silencing.

The nucleolus is the most prominent nuclear substructure and is typically surrounded by a layer of dense heterochromatin similar to the one found at the nuclear periphery. The purification of the chromatin fraction associated with nucleoli has allowed the genome-wide identification of nucleolus-associated domains (NADs) (Koningsbruggen et al., 2010; Németh et al., 2010a). Besides rDNA sequences, NADs are highly enriched in satellite repeats from pericentromeric regions. In great similarity to LADs, NADs are gene poor, transcriptionally silent and enriched in H3K27me3 and H3K9me3. Indeed, there is a significant overlap between LADs and NADs and individual heterochromatin-embedded sequences have been observed to “shuffle” between the nuclear periphery and the nucleolar surface after cell division (**Figure 14B**) (Kind et al., 2013b). Taken together, these observations suggest that the nuclear lamina and the

nucleolar surface serve as landmarks to provide similar repressive environments for heterochromatin compartmentalization (Padeken and Heun, 2014).

3.2.4. Topologically associating domains (TADs) and the loop extrusion model

One of the most notorious and conserved genome organization features revealed by HiC experiments is the segregation of chromatin into sub-megabase structures that are commonly termed “topologically associating domains” or TADs (**Figure 13D**) (Dixon et al., 2012, 2016; Sexton et al., 2012). TADs appear as triangles in the diagonal of HiC contact maps (**Figure 15A**), so that regions inside a TAD preferentially interact with regions within the same TAD, hence the concept of “topologically associating”. This implies that TADs are domains of physical chromatin contacts that are insulated from contacting adjacent regions.

High resolution HiC maps show the existence of smaller TADs, which are sometimes referred as “subTADs” (Phillips-Cremins et al., 2013; Rao et al., 2014). This suggests a nested hierarchical structure, but the field currently questions if the properties of TADs, subTADs and smaller structures are essentially different or if they are just definitions for the same features that are given depending on the resolution and computational algorithms used to identify them. Despite the wide use of the term “TAD” in the field, a unifying definition has not been reached yet.

How TAD interaction domains are created and maintained is a matter of extensive study. TAD boundaries or borders are the sites at the edges of TADs that display a high insulator capacity, preventing the regions inside the TAD from contacting with the outside of the domain. In mammals, around 40% of TADs show a distinctive frequent interaction between their two boundaries, suggesting they form a “loop”, and this type of TADs are often called “loop domains” (Rao et al., 2014). In contrast, other TADs do not appear to be created by loops but have homogeneous interaction frequencies inside the domain and are termed “compartmental domains” (Rowley et al., 2017). TAD boundaries are enriched in “insulator proteins”, specifically CTCF and the cohesin complex (Dixon et al., 2012; Rao et al., 2014). These and other observations together with computational polymer modeling approaches have led to the proposal of the “loop extrusion” model as a mechanism for loop formation and insulation that allows the creation and maintenance of TADs (**Figure 15B**) (Alipour and Marko, 2012; Fudenberg et al., 2016; Sanborn et al., 2015). In this model, the ring-shaped cohesin complex is able to “pull” chromatin through

itself, thus creating a loop by extrusion. The extrusion process is halted when cohesin encounters a barrier formed by CTCF, which interacts with cohesin only when its binding motif is found in a particular orientation, effectively resulting in a loop (Barrington et al., 2017). Although depletion of cohesin in mammalian cells leads to a loss of TAD organization, close to 20% of TAD boundaries are unaffected by removal of CTCF (Nora et al., 2017). Moreover, in *Drosophila melanogaster* there is not a loop extrusion mechanism by CTCF and cohesin but transcriptional activity seems to be a better predictor for TAD boundaries than CTCF binding (Rowley et al., 2017; Ulianov et al., 2016). These observations suggest that there are other mechanisms beside loop extrusion that must be implicated in TAD organization.

TADs tend to share a particular epigenetic “signature” with the genes inside a TAD marked by similar histone PTMs associated with either active or inactive transcription, so that most TADs overlap with domains identified in the linear genome by chromatin states (**Figure 15C**) (Dixon et al., 2012; Rao et al., 2014). This presents the TAD as a basic unit of chromatin organization to coordinate the chromatin state and gene expression of defined genomic regions, so that an entire TAD can form part of either the A or B compartment. Indeed, around 35% of the genome can switch between A and B compartments during differentiation and reprogramming with genes inside a TAD showing similar transcriptional changes, suggesting that large-scale epigenetic regulation happens at the level of TADs (Bonev et al., 2017; Dixon et al., 2015; Rao et al., 2014; Stadhouders et al., 2018).

TADs are considered to be stable and conserved between cell types, but several studies report that a fraction of TAD boundaries are different in position and insulator strength between different cell types and vary during differentiation and reprogramming (Bonev et al., 2017; Pękowska et al., 2018; Rao et al., 2014; Stadhouders et al., 2018), suggesting that these structures may be more dynamic than anticipated in some cellular processes.

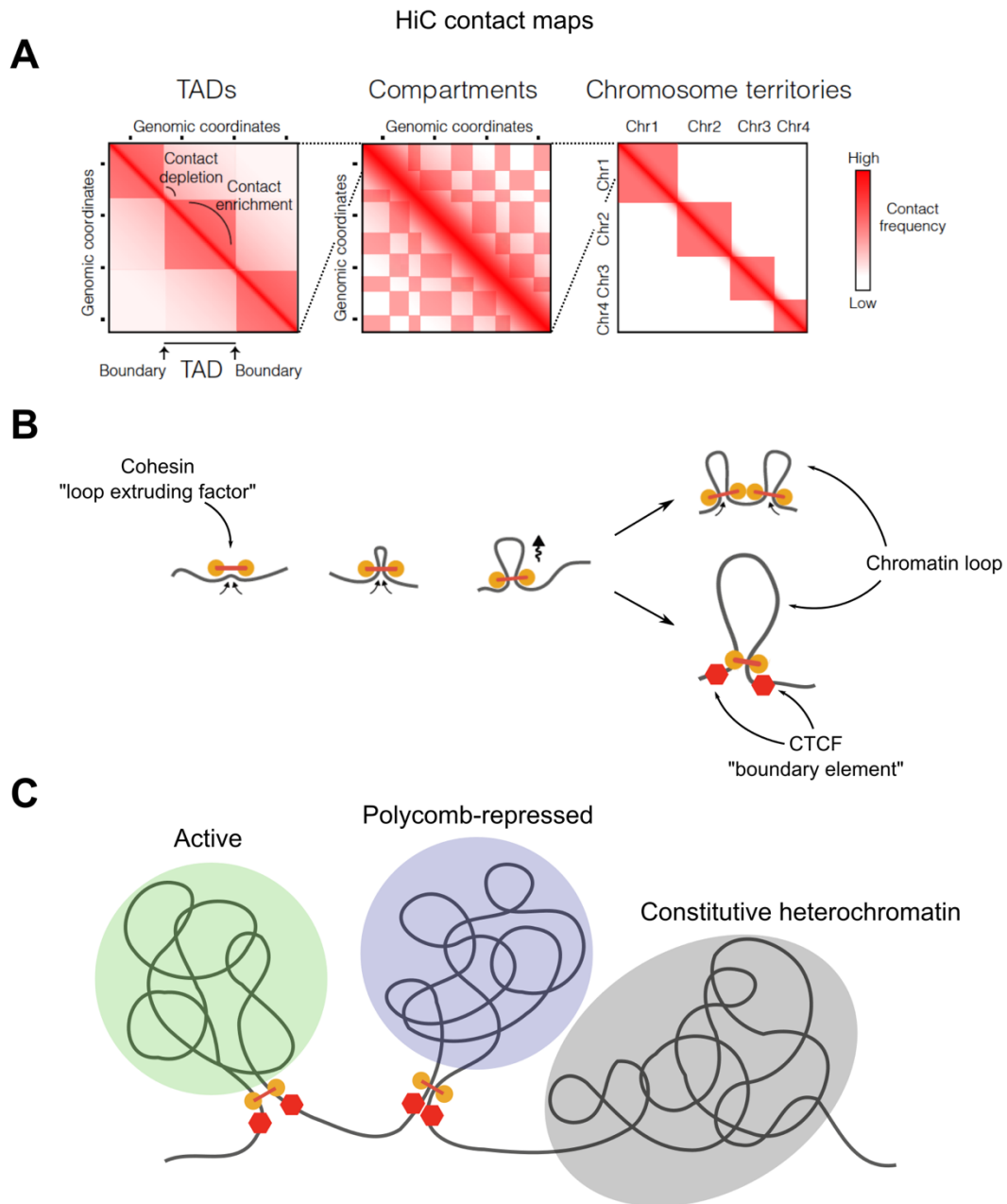


Figure 15. Topologically associated domains (TADs) and the loop extrusion model. (A) Schematic representation of HiC maps at different scales showing (from right to left) contacts between single chromosomes, compartments and TADs. The intensity of the coloring represents the frequency of interactions between two chromatin regions according to the HiC experimental data. TADs show higher interaction frequencies between the TAD while TAD boundaries show overall contact depletion. Panel adapted from Szabo et al., 2019. The HiC data is fictitious for illustrative purposes. **(B)** Schematic representation of the loop extrusion model. A loop extruding factor (cohesin) can be loaded into chromatin and begin extruding a chromatin loop until it reaches another loop extruding factor or a boundary element (CTCF). Panel adapted from Fudenberg et al., 2016. **(C)** Schematic representation of TADs sharing a common epigenetic state. Some TADs show an enrichment in CTCF and cohesin at their boundaries and have a high interaction frequency between boundaries indicative of a loop (in the picture, the "active" and "polycomb-repressed" TADs). However, other TADs show homogenous interaction and are compartmentalized independently of CTCF and cohesin (in the picture, the "constitutive heterochromatin" TAD). Panel inspired by Bonev and Cavalli, 2016.

3.2.5. Local interactions between genes and regulatory elements

Gene regulatory elements located outside of gene promoters are termed enhancers and are essential to the time and tissue-specific expression of genes that ensures proper developmental processes and cell identities (Furlong and Levine, 2018). Matching enhancer elements with their targets genes based on linear genomic data has long proven difficult, as many enhancers are found at a distance of several kilobases or even more than one megabase away from their target gene. In addition, enhancers will not always affect their nearest gene and in some cases will bypass the nearest gene to affect a farther one. Long-range promoter-enhancer contacts have been under intense study as key chromatin structures for transcriptional control.

The spatial proximity and direct contact of promoters and enhancers by chromatin looping is believed to be the main way for distal regulatory elements to control transcriptional activity (**Figure 13E**). In effect, active enhancers are generally in the proximity of active promoters (Rao et al., 2014) and forced tethering of an enhancer to a promoter in mammalian cells can induce gene activation (Deng et al., 2012). In contrast, many genes are contacting enhancers before they are activated in differentiation (Rubin et al., 2017). The interactions between the Mediator complex, specific transcription factors, CTCF and cohesin are believed to be important for the formation of promoter-enhancer loops (**Figure 16A**) (Allen and Taatjes, 2015; Stadhouders et al., 2019). In contrast to loops creating large domains, promoter-enhancers loops are often called “regulatory loops”. Although classically the contact of an enhancer with a single gene promoter was envisioned, current evidence has revealed the existence of complex structures where multiple enhancers and promoters can be found in close proximity (Bonev et al., 2017; Rao et al., 2014).

Spatial proximity is not the only factor determining the effect an enhancer has on a target gene, transcription factor binding and recruitment of co-factors activate enhancers and provides them with a characteristic chromatin state. Active enhancers are generally characterized by high DNA accessibility, an accumulation of H3K4me1 and H3K27Ac and the presence of nascent enhancer-derived RNA (eRNA) (Shlyueva et al., 2014).

Promoters and enhancers predominantly engage in contacts in the context of the same TAD (Downen et al., 2014; Schoenfelder et al., 2015; Shen et al., 2012b). This supports an organizational model where insulatory boundaries, such as TAD borders, define the effective 3D space that an enhancer can “explore”, creating confined regulatory

Introduction

landscapes for specific sets of genes (**Figure 16B**). In support of this idea, disruption of TAD boundaries results in altered promoter-enhancer interactions that can produce developmental defects (Franke et al., 2016; Guo et al., 2015; Lupiáñez et al., 2015). Moreover, a switch in the regulatory domain contacting the *HoxD* cluster determines the proper expression and function of these genes in mammalian limb development (Andrey et al., 2013).

In contrast to TADs, promoter-enhancer contacts are highly lineage and cell-type specific (Javierre et al., 2016) and appear to be very dynamic during cellular differentiation (Bonev et al., 2017). Disruption of promoter-enhancer loops or the generation of aberrant contacts with erroneous targets can lead to gene dysregulation in diseases including cancer (Krijger and de Laat, 2016). The availability of genome-wide genome conformation data and the determination of promoter-enhancer networks can allow the mechanistic understanding of the myriad of non-coding genetic variants associated with diseases (Javierre et al., 2016; Krijger and de Laat, 2016).

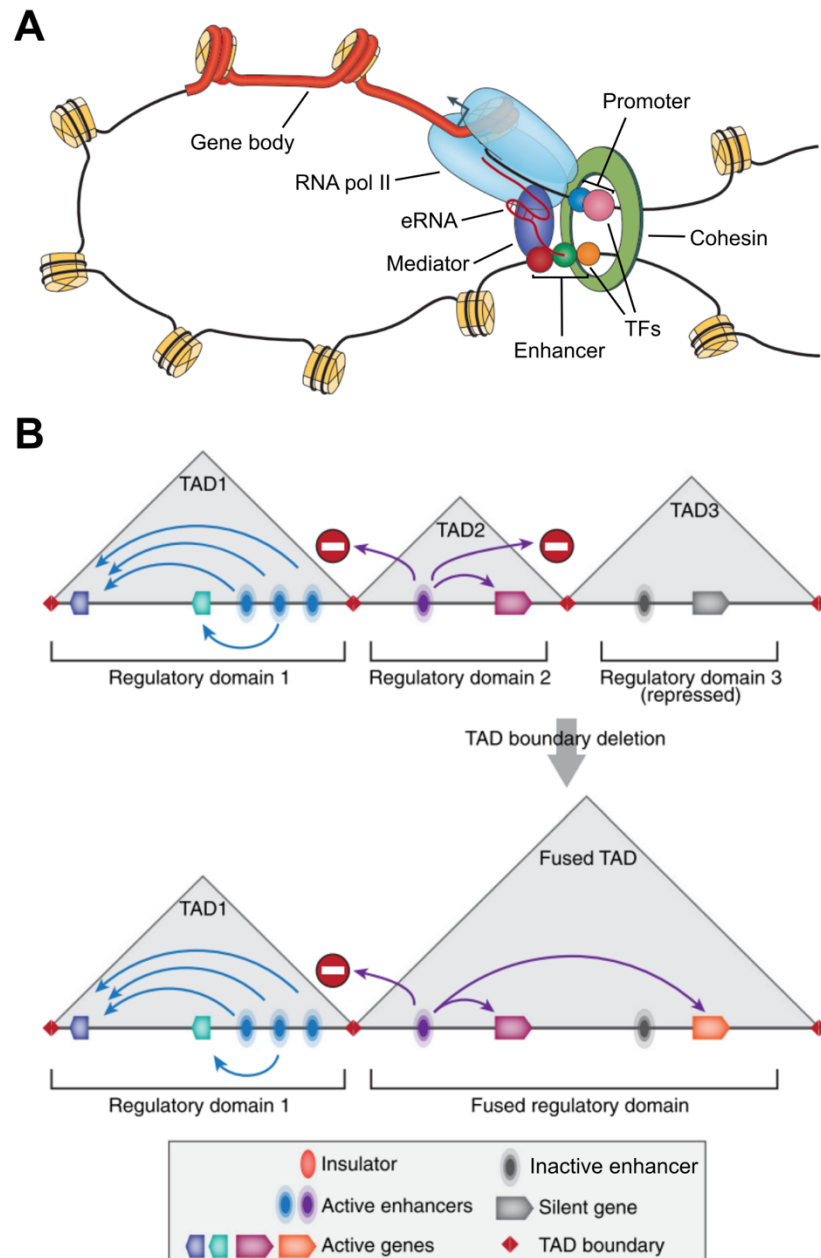


Figure 16. Promoter-enhancer contacts. **(A)** Schematic model of the factors involved in the formation of contacts between promoters and enhancers. Chromatin loops bring promoters and enhancers bound by transcription factors in close proximity. This interaction is thought to be stabilized by the Mediator complex, the cohesin ring complex and in some cases by enhancer-derived RNAs (eRNAs). Panel adapted from Bonev and Cavalli, 2016. **(B)** Schematic representation of the current model where enhancer interaction is mainly limited to the TAD environment. TAD boundaries create regulatory domains that limit the influence of enhancers to genes within the domain. Disruption of TAD boundaries can “fuse” regulatory domains and permit aberrant interactions between active enhancers and promoters leading to unwanted gene expression.

Aims and Objectives

4. Aims and objectives

Several studies on development, differentiation, cancer and somatic cell reprogramming have suggested that macroH2A histone variants have a role as epigenetic stabilizers of differentiated states. However, the molecular mechanisms underlying this function remain largely unknown. In particular, although macroH2A is mostly related to transcriptionally silenced states, its reported functions in gene are highly context-dependent and include repression and activation events. The aim of my PhD has been to characterize the role of the group of macroH2A histone variants in chromatin organization and transcription with the following specific objectives:

1. Study the role of macroH2A in the nuclear organization of chromatin.
2. Characterize the molecular mechanisms and domain requirements of the function of macroH2A in chromatin organization.
3. Describe the macroH2A-dependent changes in the transcription and phenotype of cells in a cancer model.

Results

5. Preamble

The results obtained during the development of my doctoral thesis have been included in two published articles and one manuscript currently in preparation. The results chapter is subdivided in three sections that separate the results in blocks related to these three manuscripts. My relevant and increasing contributions to these articles are reflected in the fact that I am signing them as second author of 13 authors (Douet et al., 2017), as shared first author (Kozlowski et al., 2018) and sole first author (see below). I have divided the results section of my thesis in three chapters that essentially follow the three manuscripts. For completeness and with the intention to be able to coherently discuss the results of the projects as a whole, I have included results that were obtained by other team members, collaborators or that we accomplished through shared collaborative efforts. At the end of each results section I include a paragraph clarifying my own contributions and detailing those by others.

Douet J, **Corujo D**, Malinverni R, Renauld J, Sansoni V, Posavec Marjanović M, Cantariño N, Valero V, Mongelard F, Bouvet P, Imhof A, Thiry M, Buschbeck M (2017). MacroH2A histone variants maintain nuclear organization and heterochromatin architecture. *J. Cell Sci.*

Corujo, D.*, Kozlowski, M.*, Hothorn M, Guberovic I, Mandemaker IK, Blessing C, Sporn J, Gutierrez-Triana A, Smith R, Portmann T, Treier M, Scheffzek K, Huet S, Timinszky G, Ladurner AG**, Buschbeck M** (2018). MacroH2A histone variants limit chromatin plasticity through two distinct mechanisms. *EMBO Rep.* *, Shared first-authorship, inverted print order. **, shared corresponding authorship, inverted print order.

Corujo D et al., and M. Buschbeck (2019) MacroH2A represses the expression of DKK1 and attenuates its transcriptional activation by TNF α (Working Title). Manuscript in preparation.

6. Results I: MacroH2A maintains nuclear organization and heterochromatin architecture

6.1. MacroH2A is essential for maintaining nuclear organization

In order to study the role of macroH2A in global nuclear organization, we depleted the expression of both macroH2A1 and macroH2A2 proteins in human cells. We selected the human male cell line HepG2 as a model system. HepG2 is a well-characterized hepatoblastoma cell line, is included in the ENCODE (ENCODE Project Consortium et al., 2013) project and has been used to generate many genomic datasets. HepG2 cells express both macroH2A1 and macroH2A2 proteins (**Figure 17A**). We used stable retroviral integration of shRNA cassettes to achieve a highly-efficient double knockdown of both macroH2A1 and macroH2A2 protein levels (**Figure 17A, B**). We have named the cells with depleted macroH2A **DKD** (which stands for double knockdown) and the corresponding control cells **RDM** (which stands for the random sequence shRNA used as control).

Transmission electron microscopy images of the cells' nuclei show that depletion of macroH2A proteins had a drastic effect on the shape of the nucleus and its internal organization (**Figure 17C**). Control cells mostly have a rounded nucleus with regions of dense heterochromatin staining throughout the nucleoplasm, particularly concentrated at the nucleolar surface and the nuclear periphery (**Figure 17C**). DKD cells exhibit a global loss of heterochromatic staining, particularly noticeable at the nucleolar and nuclear periphery (**Figure 17C, D**). This phenotype had a high penetration when analyzing 100 cells of each group (**Figure 17E**). Additionally, nucleoli were expanded in DKD cells with a reticulated pattern and several fibrillar centers (**Figure 17D**). These nucleolar structural changes are associated with increased transcription of ribosomal DNA (rDNA), a previously reported effect of depleting macroH2A1 in HepG2 cells (Cong et al., 2014).

Results

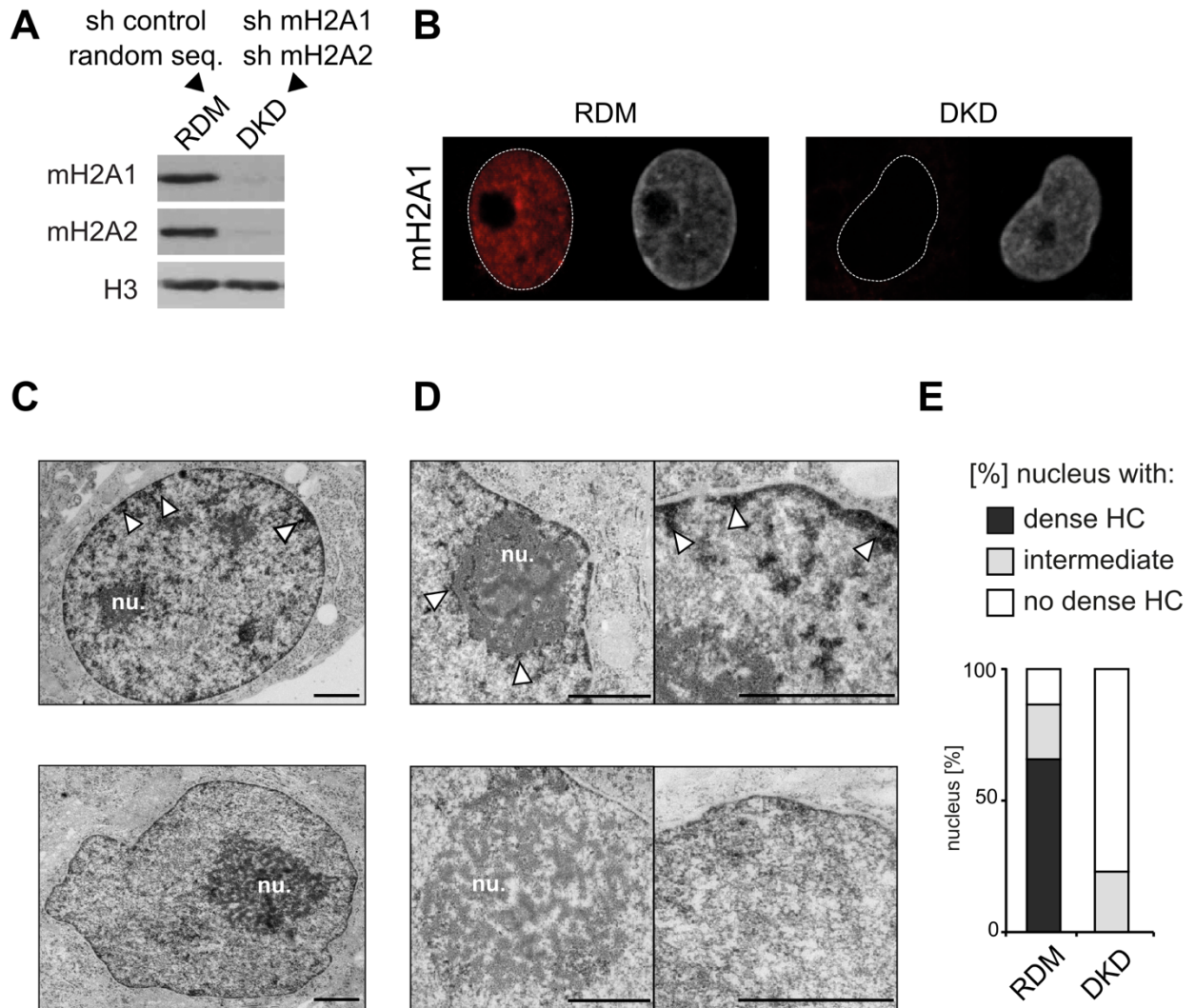


Figure 17. MacroH2A loss impairs global nuclear organization. **(A)** Western blot showing the reduction in protein levels of macroH2A1 and macroH2A2 in HepG2 cells achieved by stable integration of two shRNAs targeting macroH2A1 and macroH2A2 separately (DKD) in comparison with control cells transduced with a random sequence shRNA (RDM). H3 is included as loading control. **(B)** Immunofluorescence of macroH2A1 (red) in RDM and DKD cells. Nuclei were counterstained with DAPI (grey). **(C)** Transmission electron microscopy images of nuclei from RDM and DKD cells. White arrowheads indicate areas of dark staining corresponding to heterochromatin and "nu." marks the nucleolus. Scale bars are 1 μ . **(D)** Higher magnification images focused on the nucleolus and the nuclear periphery. White arrowheads indicate areas of dark staining corresponding to heterochromatin and "nu." marks the nucleolus. Scale bars are 1 μ . **(E)** Quantification of nuclei as imaged in C according to three classes based on their heterochromatin content (HC) (n> 100).

We further confirmed the nucleolar expansion by performing immunofluorescence detection of nucleophosmin (NPM1) as a nucleolar marker (**Figure 18A**). Automated measurement of nuclei and nucleoli on 2D projections of confocal image stacks showed an increase in the proportion of the nuclear area corresponding to the nucleolus as well as an overall enlargement of nuclei in DKD cells (**Figure 18B, C**). The transmission electron microscopy images showed that some nuclei have an irregular shape (**Figure 17C**). By visual inspection of nuclei shape based on DAPI staining we observed an increase in the proportion of nuclei with anomalies (such as buds or cavities) in DKD cells (**Figure 18D, E**).

Taken together, these results highlight an important role of macroH2A in the maintenance of the proper nuclear organization of heterochromatic and nucleolar structures.

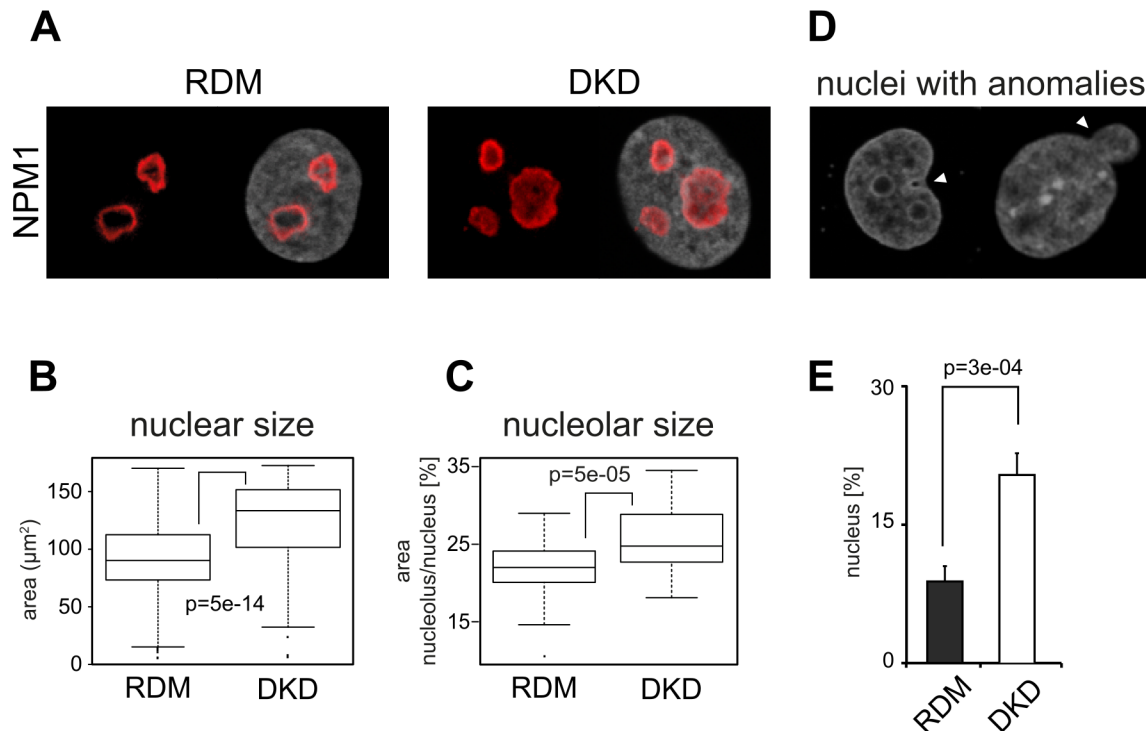


Figure 18. MacroH2A loss results in larger nuclei and nucleoli and an increased appearance of nuclear anomalies. **(A)** Confocal immunofluorescence of nucleophosmin (NPM1, red) in RDM and DKD cells. Nuclei were counterstained with DAPI (grey). **(B)** Quantification of nuclear size measured as the DAPI area on a 2D projection of confocal images of RDM and DKD cells stained as in A. The pvalue (p) is calculated by a Wilcoxon rank sum test ($n > 100$). In the boxplots, the box represents the 25-75th quartiles and the median is indicated as a horizontal line. The whiskers cover the highest and lowest data points within the 1.5x interquartile range of the upper and lower quartile, respectively, and outliers appear as points. **(C)** Percentage of nuclear area taken up by all the nucleoli in a given cell (as defined by NPM1 staining) measured as in B. The pvalue (p) is the result of a Wilcoxon rank sum test (RDM $n = 49$, DKD $n = 51$). **(D)** Example confocal images of anomalies in nuclei shape observed in HepG2 cells. **(E)** Quantification by visual inspection of the proportion of nuclei with anomalies in RDM and DKD cells. The pvalue (p) is calculated by a two-tailed Z-test. ($n > 200$, three biological replicates).

6.2. A fraction of macroH2A is associated with H3K9me3 and repetitive elements

To understand the relationship between macroH2A and the heterochromatic compartment, we performed ChIP-Seq experiments with antibodies targeting separately macroH2A1 and macroH2A2. The anti-macroH2A1 antibody used was raised against the extranucleosomal part of the protein and does distinguish the two splice variants of macroH2A1 (see Materials and Methods). To study the link between macroH2A enrichment and heterochromatin, we used available ENCODE datasets with ChIP-Seq data for H3K27me3 and H3K9me3, two major marks of facultative and constitutive heterochromatin, respectively, as well as RNA-Seq expression data. The analysis of the genomic distribution of macroH2A according to our ChIP-Seq data shows several characteristics that have been previously reported in other cell types (**Figure 19A**). In general, macroH2A is found enriched in large domains up to hundreds of kb (Gamble et al., 2010) and there is a high similarity between macroH2A1 and macroH2A2 (Gaspar-Maia et al., 2013; Pehrson et al., 2014). Given this similarity, we focused our subsequent analysis on macroH2A2. MacroH2A2 occupancy at genic regions correlates with low transcription (Changolkar and Pehrson, 2006; Changolkar et al., 2010; Gamble et al., 2010), while a large fraction of macroH2A2 is associated with the presence of H3K27me3 (Chen et al., 2014; Gamble and Kraus, 2010; Gaspar-Maia et al., 2013).

Interestingly, a smaller fraction of macroH2A2 bound chromatin is associated with H3K9me3 (**Figure 19B**). Of the total of 1670 macroH2A2 peaks that overlap with H3K9me3, 1180 do so without overlapping in turn with H3K27me3. These 1180 peaks represent around 10% of the total macroH2A2 called peaks and consist of 1 to 5kb regions isolated in euchromatic environments. In contrast to the more common distribution of macroH2A in the form of large domains, these peaks are easily identified as narrow peaks of high enrichment for macroH2A1, macroH2A2 and H3K9me3.

As the hallmark of constitutive heterochromatin, H3K9me3 is found in many repetitive elements and is an essential factor in their silencing. We questioned if macroH2A could be associated with H3K9me3 in repetitive regions that had been initially excluded from our ChIP-Seq mapping and peak calling. Analysis of genomic repeats from high-throughput sequencing data poses a bioinformatic challenge, and standard ChIP-Seq analysis pipelines discard most repetitive sequence reads as they cannot be uniquely mapped in the genome. Still, through the use of Repeatmasker, a fraction of repetitive

elements is annotated in the genome build. We have used the R package *Regioner* (Gel et al., 2016) to perform permutation-based association tests between macroH2A peaks and different classes of annotated repeats. We found that macroH2A1 and macroH2A2 peaks are positively associated with different classes of repeats (**Figure 19C**). Analyzing the fraction of macroH2A2 peaks that coincides with H3K9me3, most of the associations become statistically stronger. In particular, macroH2A2-H3K9me3 peaks show significant positive associations with simple repeats, LINEs, LTRs, satellites, ribosomal RNA and small nuclear RNAs (**Figure 19C**).

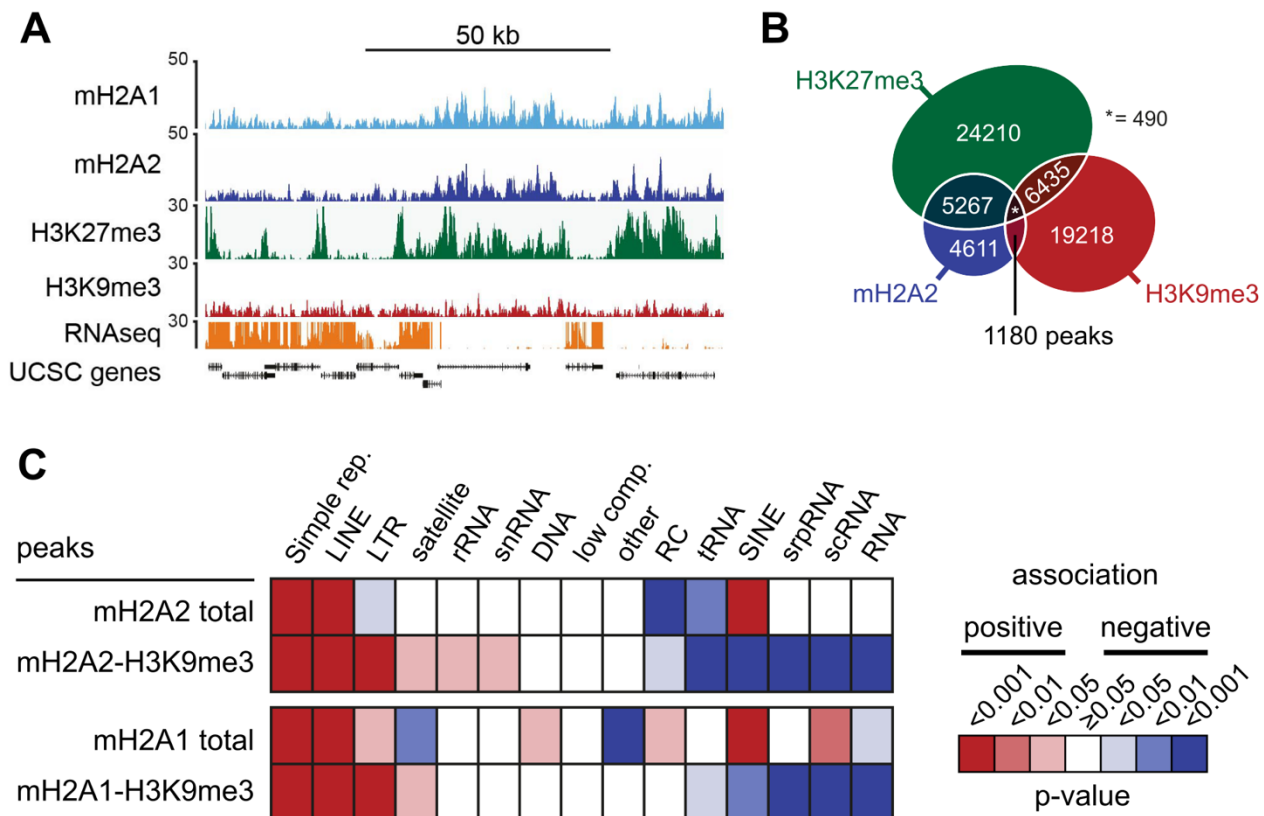


Figure 19. MacroH2A associates with H3K9me3 and repetitive elements. (A) Screenshot from the UCSC genome browser showing the ChIP-Seq profiles of macroH2A1 and macroH2A2 and RNA-Seq, H3K27me3 and H3K9me3 from ENCODE (B) Venn diagram showing the overlap between macroH2A2 peaks with H3K9me3 and H3K27me3 peaks in HepG2 cells. (C) Heatmap summarizing the association of macroH2A2 peaks and regions containing macroH2A2 and H3K9me3 with the repetitive elements indexed in Repeatmasker. The association values were calculated with the permutation test-based R package *regioner* (Gel et al., 2016) performing 1000 iterations. Red and blue colors denote positive and negative associations, respectively, and the color intensity relates to the p-value.

Following this analysis, we used ChIP-qPCR to test the enrichment of macroH2A in large heterochromatic repetitive regions that are still mostly excluded from genomic assemblies. These repeats included the pericentromeric satellite *SAT2*, the centromeric α -satellite, 45S and 5S rDNA and *DXZ4*, a repeat present in the X chromosome and marked by H3K9me3 in male cells (Chadwick, 2008). We found a significant enrichment signal for macroH2A1 and macroH2A2 in all the repeats tested that decreased close to

Results

background levels in DKD cells, demonstrating that the signal is specific for the presence of macroH2A (**Figure 20A, B**). Knockdown of macroH2A did not affect the presence of H3K9me3 in the tested repetitive elements (**Figure 20C**). Similar observations were made in mouse embryonic fibroblasts (MEF) from control mice and macroH2A1 knock-out (*mH2A1* ^{-/-}) (**Figure 20D**). MacroH2A1 is enriched in Line-L1, minor and major satellites and rDNA, and its loss did not affect the presence of H3K9me3 on these repetitive elements (**Figure 20E, F**).

These results demonstrate that macroH2A proteins are enriched in genomic repetitive elements that form part of the H3K9me3 marked constitutive heterochromatin.

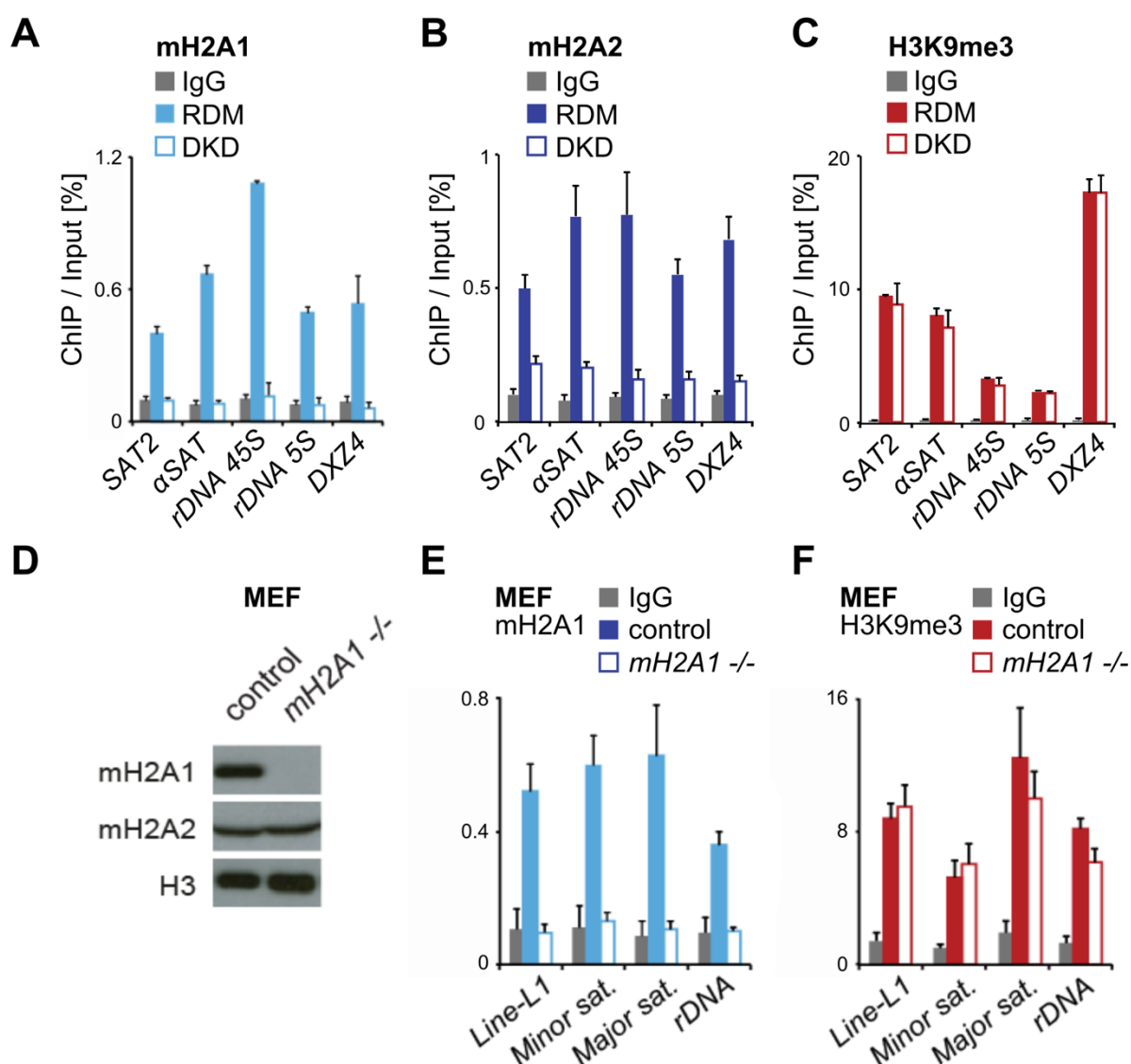


Figure 20. MacroH2A is enriched at repetitive elements but is not necessary for H3K9me3 marking. (A, B, C) ChIP-qPCR showing the enrichment of macroH2A1, macroH2A2 and H3K9me3 at certain repetitive elements in HepG2 RDM and DKD cells. (D) Western blot showing the levels of macroH2A1 and macroH2A2 in mouse embryonic fibroblasts (MEF) from control and *mH2A1* ^{-/-} mice. H3 is included as a loading control. (E, F) ChIP-qPCR showing the enrichment of macroH2A1 and H3K9me3 at certain repetitive elements in control and macroH2A1 ^{-/-} MEFs.

6.3. MacroH2A maintains the architecture of heterochromatic repetitive elements

To investigate which function macroH2A could have on the repeats where it is enriched, we used Fluorescence In Situ Hybridization (FISH) to look at the organization of two particular elements: the pericentromeric satellite *SAT2* and 45S rDNA. In the human genome, there are approximately 300 copies of 45S rDNA distributed in five clusters on the short arms of five acrocentric chromosomes (Diesch et al., 2014). We observed an increased number of rDNA foci in DKD cells when compared to RDM, indicative of a disorganization of the rDNA sequences (**Figure 21A, B, top panels**). Confocal imaging of a FISH probe targeting the human *SAT2* repeat located in chromosome 1 showed significant alterations in DKD cells. While in control cells *SAT2* is detected as a small spot-like signal, loss of macroH2A provoked the appearance of expanded and fibre-like structures in many cells which is quantified as an overall greater area of *SAT2* signal in 2D projections (**Figure 21A, B, middle panels**). We used H3K9me3 immunostaining to analyze globally constitutive heterochromatin and observed a general diffusion of the signal in DKD cells, with less well defined high-intensity H3K9me3 foci than in RDM cells (**Figure 21A, bottom panel**). This was quantified as a smaller total H3K9me3 area in 2D projections when subjecting the images to an automated thresholding aimed at detecting bright H3K9me3 foci (**Figure 21B, bottom panel**).

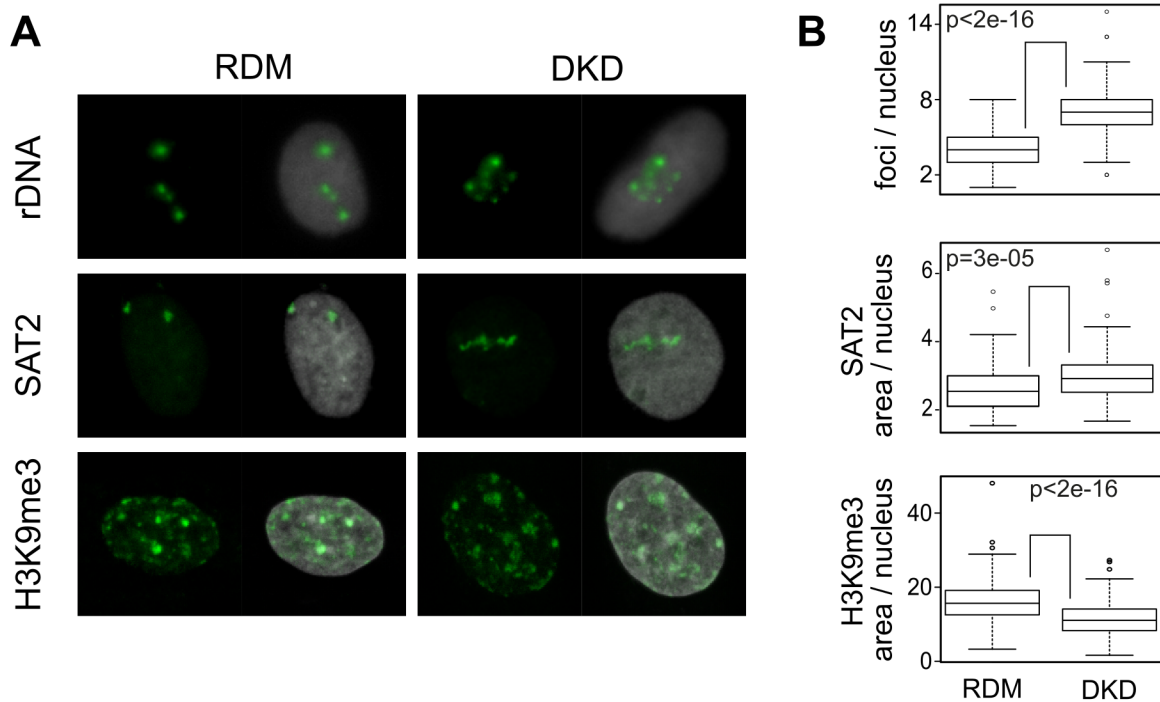


Figure 21. MacroH2A maintains the architecture of heterochromatic elements. (A) Fluorescence In Situ Hybridization (FISH) of 45S rDNA and SAT2 (top and middle panels) and H3K9me3 immunostaining (bottom panel) in RDM and DKD cells. Nuclei were counterstained with DAPI (grey). **(B)** Quantification of the images in (A). Top panel: number of rDNA foci per nucleus (RDM $n = 81$, DKD $n = 62$). Middle panel: percentage of nuclear area corresponding to SAT2 FISH signal (RDM $n = 79$, DKD $n = 125$). Bottom panel: percentage of nuclear area corresponding to bright H3K9me3 foci (RDM $n = 229$, DKD $n = 215$). The p-value (p) is calculated by a Wilcoxon rank sum test. In the boxplots, the box represents the 25-75th quartiles and the median is indicated as a horizontal line. The whiskers cover the highest and lowest data points within the 1.5x interquartile range of the upper and lower quartile, respectively, and outliers appear as points.

Next, we wondered if the structural alterations in heterochromatic repetitive elements observed upon macroH2A loss would be accompanied by changes in the compaction of chromatin at the level of the DNA fiber. We used two methods to assess chromatin accessibility and nucleosomal density. First, we used a Formaldehyde Assisted Isolation of Regulatory Elements (FAIRE) assay (Simon et al., 2012), designed to test the accessibility or openness of chromatin regions and initially aimed at identifying nucleosome-depleted regulatory elements. In our case, although we could observe different FAIRE signals for characterized “closed” and “open” loci, we did not observe changes in the openness of chromatin at repetitive elements in the absence of macroH2A (**Figure 22A**). Second, we used ChIP-qPCR targeting histone H3 to test the nucleosomal occupancy of the same regions and did not detect any difference between control and DKD cells (**Figure 22B**). Note that the signal profiles across the different tested regions are inverted between FAIRE and H3 ChIP (that is, a region with a low openness FAIRE signal shows a high H3 ChIP nucleosomal density signal, and vice-versa), which indicates

that these two complementary approaches are effective to assess chromatin compaction at the DNA fiber level. To test whether the loss of macroH2A affected the transcriptional activity of repeats we used reverse transcription quantitative PCR (RT-qPCR). The levels of transcripts for the tested repeats were increased between 1.5 and 3-fold in DKD cells compared to control cells (**Figure 22C**).

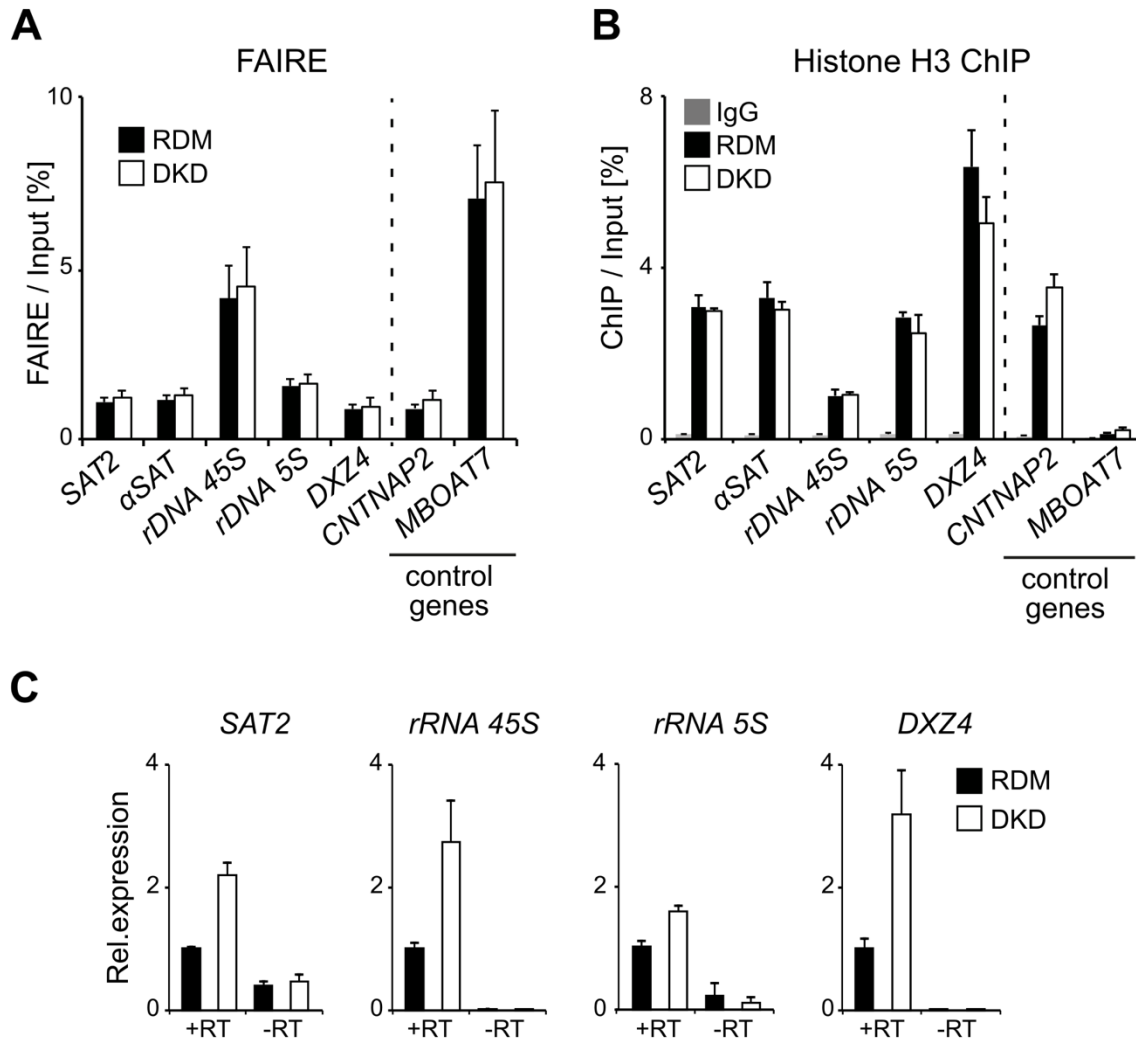


Figure 22. MacroH2A loss does not affect nucleosomal density of repetitive elements and mildly increases their transcription. (A) Local chromatin openness of repetitive elements assayed by FAIRE. The FAIRE signal value is calculated as the percentage of FAIRE-extracted DNA with respect to input DNA. The genes CNTNAP2 and MBOAT7 are included as control loci previously described to be closed and open, respectively (Simon et al., 2012). Data is plotted as the mean and error bars represent the SEM (n = 3). (B) H3 occupancy at repetitive elements assayed by ChIP. The same control loci as in (B) are used. Data is plotted as the mean and error bars represent the SEM (n = 3). (C) Relative expression of RNA from four different repetitive elements determined by RT-qPCR. Negative control reactions without reverse transcriptase (-RT) were included to determine the technical background signal. Values are normalized to the endogenous control genes GAPDH and RPLP0. Data is plotted as the mean and error bars represent the SEM (n = 3).

Results

These results show that the structural alterations of repetitive elements upon loss of macroH2A are not a consequence of a loss of H3K9me3 or condensation of the chromatin fiber and have a low impact in their transcription. Taken together, these observations suggest that such alterations occur at the level of higher-order chromatin organization.

6.4. MacroH2A promotes the attachment of repetitive elements to lamin B1 filaments

The nuclear periphery, together with the nucleolar surface, accumulates the major fraction of heterochromatin in the cell and is considered as an important repressive compartment in particular for the repetitive elements that form constitutive heterochromatin. As the loss of heterochromatic staining in DKD cells was particularly evident at the nuclear periphery, we wondered whether macroH2A could participate in the anchoring of repetitive elements to the nuclear lamina. The knock-down of lamin B1, an integral protein of the filament network that forms the nuclear lamina, resulted in alterations of the structure of the *SAT2* repeat that were very similar to those observed in cells lacking macroH2A (**Figure 23A-C**).

Thus, we interrogated if macroH2A-bound chromatin is interacting with lamin B1 at the nuclear periphery. For this, we used an *in situ* Proximity Ligation Assay (PLA) (Söderberg et al., 2006), which is based on the proximity of two antibodies detecting different epitopes allowing the ligation and amplification of a DNA template that can then be detected with fluorescent oligonucleotide probes (**Figure 24A**). PLA can detect interactions up to 40nm in distance and in the context of the 11 to 30nm chromatin fibers has proven to be a useful technique for the identification of interactions between histone modifications, chromatin regulators and lamin proteins (Barateau and Buendia, 2010; Cesarini et al., 2015). Using PLA, we could detect interactions of both macroH2A1 and macroH2A2 with lamin B1 (**Figure 24B, C**). The interactions are found mostly in the nuclear periphery and are greatly reduced in cells lacking macroH2A, demonstrating the

specificity of the signal. We could also detect interactions between macroH2A1 and H3K9me3 which concentrated mostly in the nuclear periphery (**Figure 24D**).

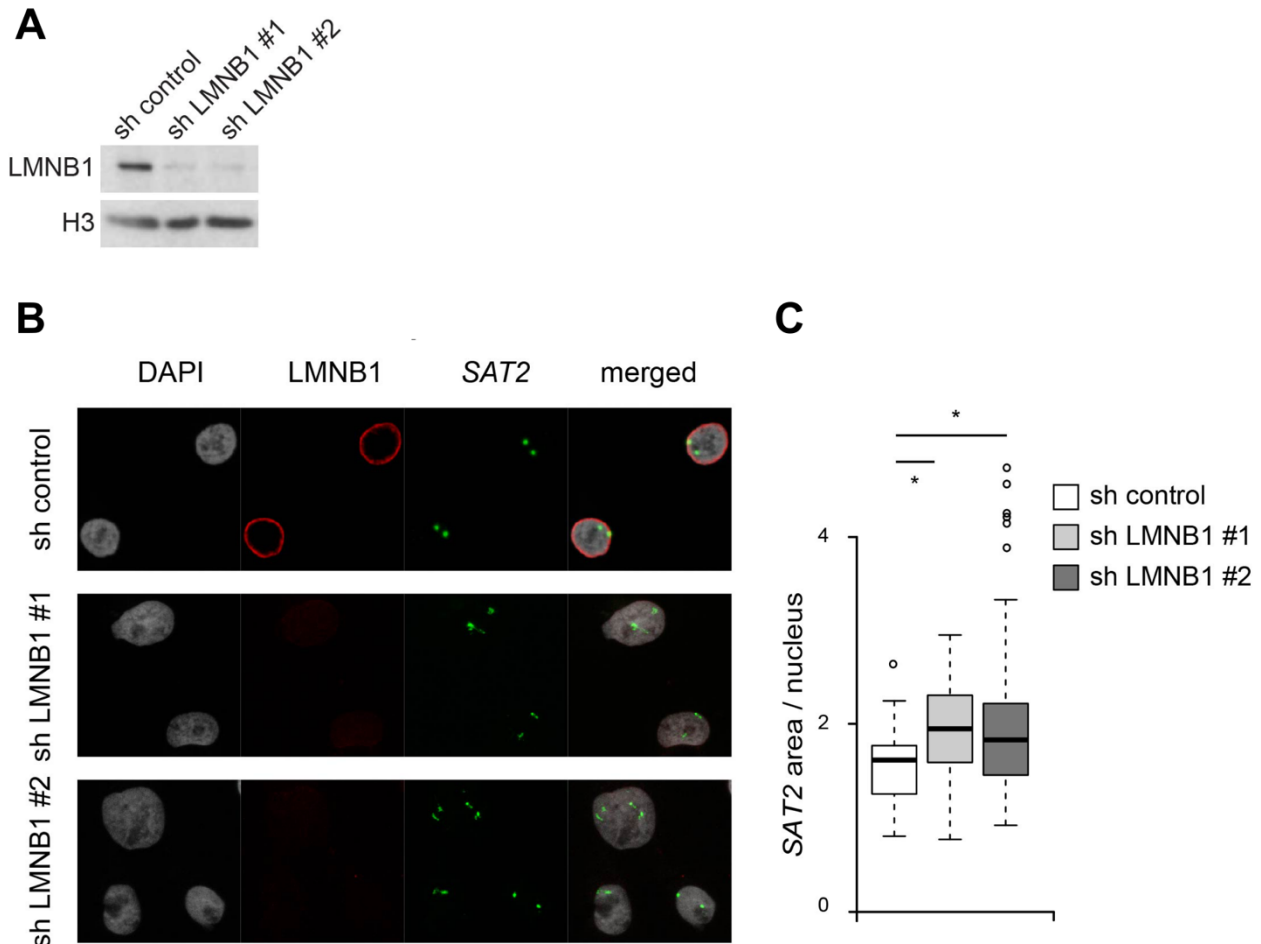


Figure 23. Alterations of repetitive elements caused by lamin B1 knockdown are similar to those caused by the absence of macroH2A. (A) Western blot showing the expression of lamin B1 (LMNB1) in HepG2 cells transduced with either a control shRNA (sh control) or two independent shRNAs targeting lamin B1 (sh LMNB1 #1 and sh LMNB1 #2). H3 is included as a loading control. (B) Immuno-FISH combining the simultaneous detection of LMNB1 (red) and SAT2 repeats (green) in control and lamin B1 knockdown cells. Nuclei were counterstained with DAPI (grey). (C) Quantification of the percentage of nuclear area corresponding to SAT2 FISH signal. A Wilcoxon rank sum test was used to compare the sh control group with the two sh LMNB1, * $p < 0.05$ (sh control $n = 52$, sh LMNB1 #1 $n = 58$, sh LMNB1 #2 $n = 106$).

Results

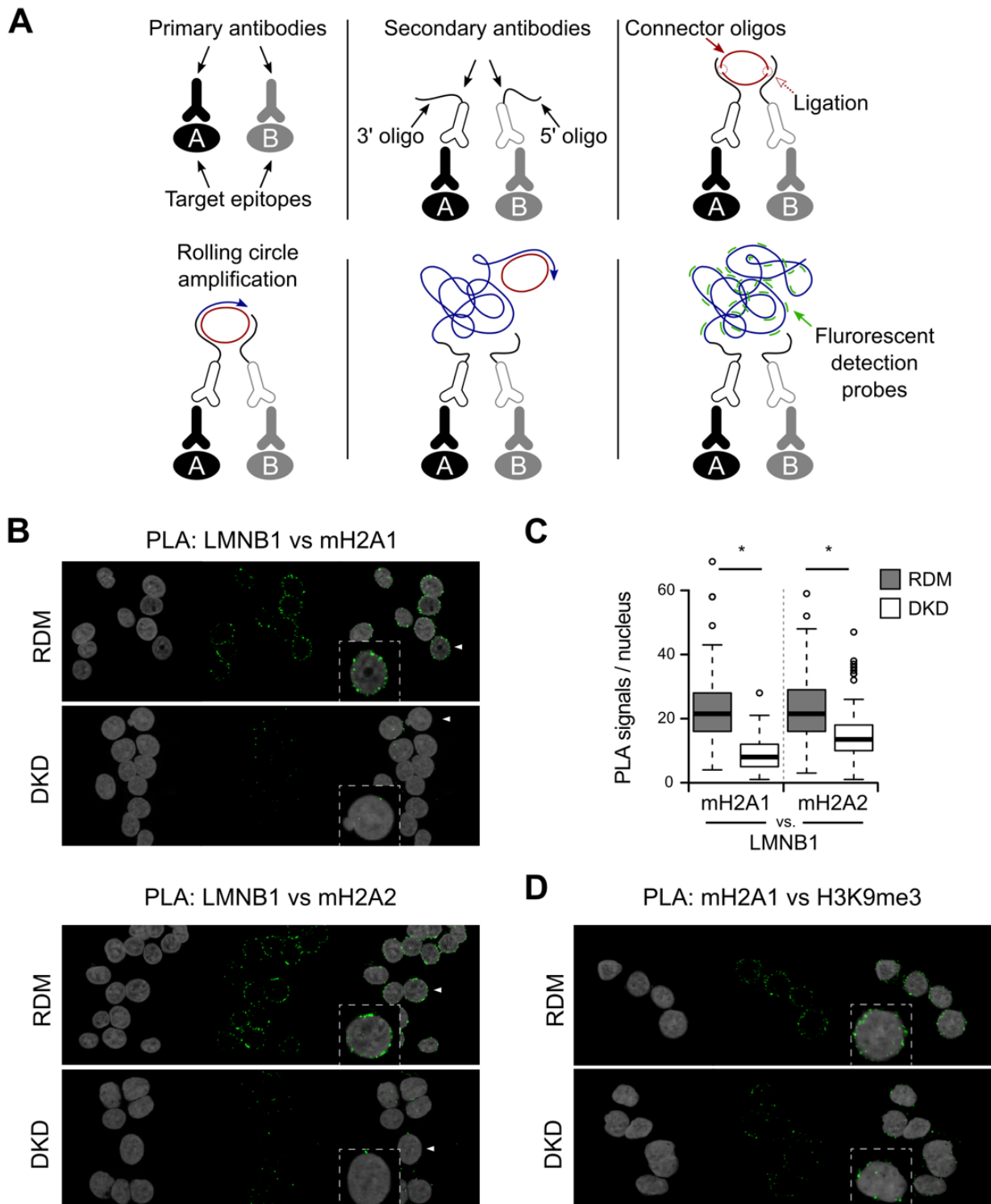


Figure 24. MacroH2A proteins interact with lamin B1. (A) Schematic representation of the Proximity Ligation Assay (PLA). Two primary antibodies recognize specific epitopes of interest. In turn, secondary antibodies coupled with oligonucleotides detect the primary antibodies. A connector oligo is added and, if the secondary antibodies are 40nm or less apart, a ligation event produces a circular ssDNA. This template can be amplified by a DNA polymerase in a rolling circle amplification creating a large repetitive ssDNA molecule bound to the interaction site. Finally, hybridization of fluorescence-coupled oligos allows the detection of the PLA signals as discrete spots under the microscope. (B) Detection of macroH2A1 or macroH2A2 and lamin B1 (LMNB1) interactions by PLA (green). Nuclei were counterstained with DAPI (grey). The arrowhead indicates the nucleus in the outlined zoom-in. (C) Quantification of the number of PLA per nucleus in the images in (A). A Wilcoxon rank sum test was used to compare the RDM and DKD groups, * $p < 0.05$ ($n \geq 100$). (D) Detection of macroH2A1 and H3K9me3 interactions by PLA (green). Nuclei were counterstained with DAPI (grey). The arrowhead indicates the nucleus in the outlined zoom-in.

Next, we asked if the interaction of constitutive heterochromatin and repetitive elements with lamin B1 would require the presence of macroH2A. First, we checked that lamin B1 protein levels and nuclear localization were not affected in the absence of macroH2A (**Figure 25A, B**). Then, we performed PLA experiments to test the interaction of H3K9me3 with lamin B1. As expected from the nuclear localization of lamin B1, most PLA signals were observed at the nuclear periphery (**Figure 25C**). We observed a reduction of around 40% in the detected interactions between H3K9me3 and lamin B1 in DKD cells compared to control cells (**Figure 25C, D**). Finally, we performed ChIP experiments to determine if the loss of macroH2A affected the interaction of specific repetitive sequences with lamin B1. We found lamin B1 enriched in all the tested repeats and the loss of this enrichment in the absence of macroH2A (**Figure 26A**). To validate this finding, we performed the same experiment in mESC cell lines transduced with either a control or macroH2A-targeting shRNA. We achieved an efficient depletion of macroH2A protein levels in mESC and did not find lamin B1 protein levels to be affected (**Figure 26B**). Similarly to the observation made in HepG2, ChIP for lamin B1 found the repeats tested enriched in lamin B1 and lost the signal upon macroH2A1 knockdown (**Figure 26C**).

Taken together, these results demonstrate a function for macroH2A in nuclear organization. This function includes the maintenance of heterochromatin architecture mediated at least in part by promoting the interaction of H3K9me3 and repetitive elements with lamin B1.

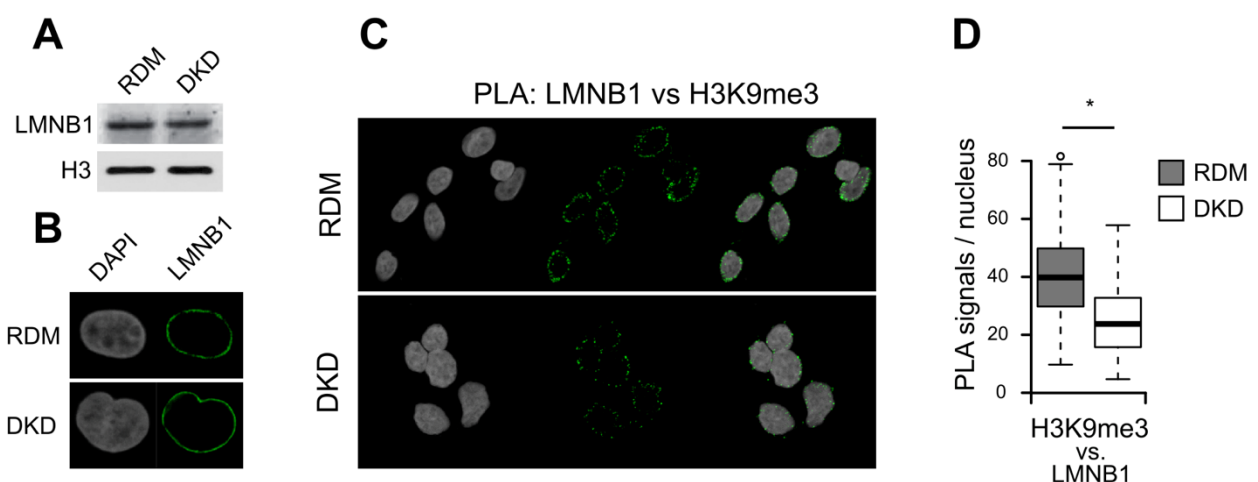


Figure 25. The absence of macroH2A reduces the interactions between H3K9me3 and lamin B1. (A) Western blot showing the expression of lamin B1 in RDM and DKD cells. H3 is included as a loading control. (B) Immunofluorescence detection of lamin B1 (green) in RDM and DKD cells. Nuclei were counterstained with DAPI (grey). (C) Detection of lamin B1 and H3K9me3 interactions by PLA (green). Nuclei were counterstained with DAPI (grey). (D) Quantification of the number of PLA signals per nucleus in the images in (C). A Wilcoxon rank sum test was used to compare the RDM and DKD groups, * $p < 0.05$ ($n > 97$).

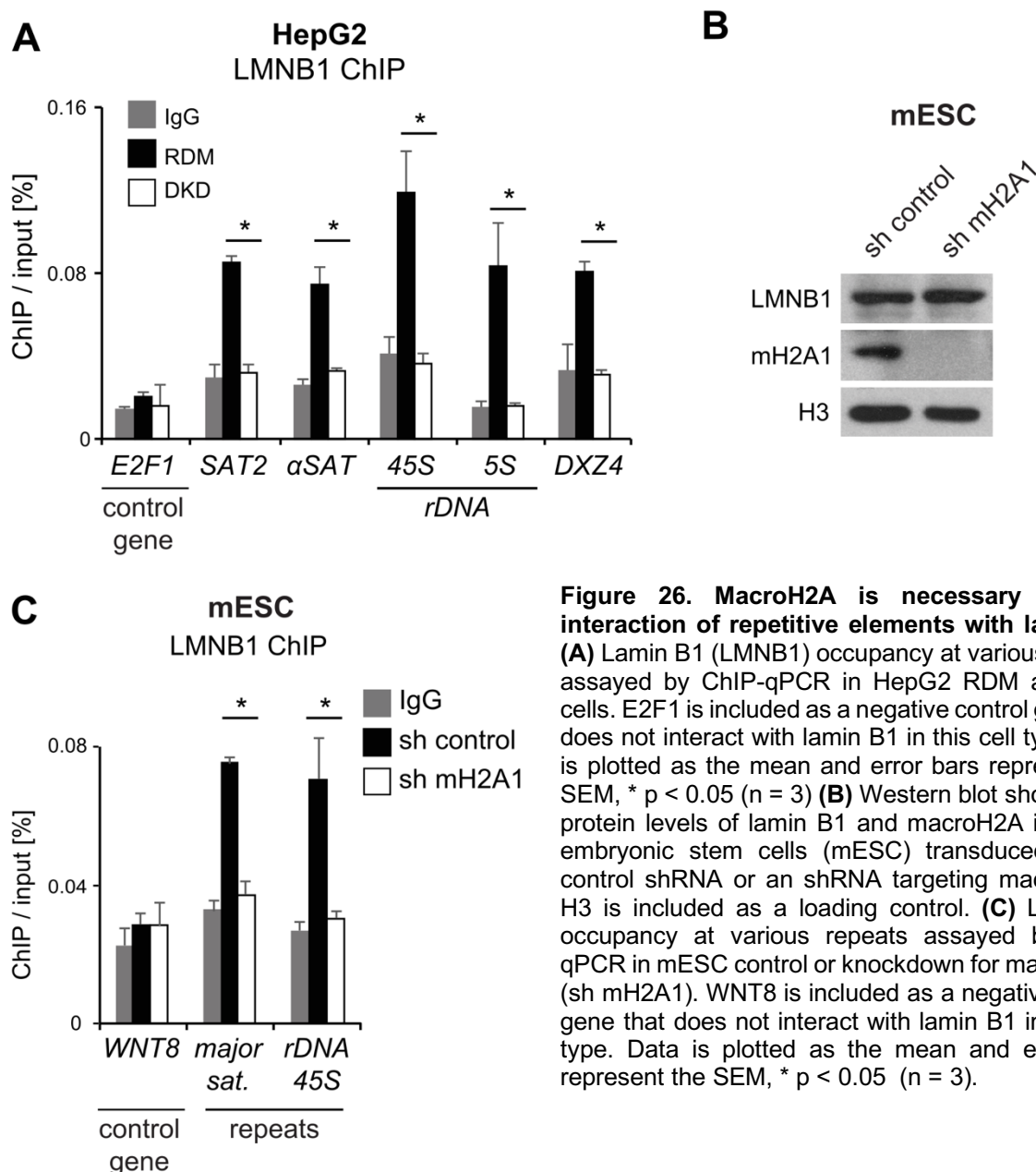


Figure 26. MacroH2A is necessary for the interaction of repetitive elements with lamin B1.

(A) Lamin B1 (LMNB1) occupancy at various repeats assayed by ChIP-qPCR in HepG2 RDM and DKD cells. E2F1 is included as a negative control gene that does not interact with lamin B1 in this cell type. Data is plotted as the mean and error bars represent the SEM, * $p < 0.05$ ($n = 3$) **(B)** Western blot showing the protein levels of lamin B1 and macroH2A in mouse embryonic stem cells (mESC) transduced with a control shRNA or an shRNA targeting macroH2A1. H3 is included as a loading control. **(C)** Lamin B1 occupancy at various repeats assayed by ChIP-qPCR in mESC control or knockdown for macroH2A1 (sh mH2A1). WNT8 is included as a negative control gene that does not interact with lamin B1 in this cell type. Data is plotted as the mean and error bars represent the SEM, * $p < 0.05$ ($n = 3$).

6.5. Contributions

This project was conceived and supervised by Julien Douet and Marcus Buschbeck. Julien Douet generated the HepG2 RDM and DKD cells as well as the LMNB1 knock-downs, performed and imaged FISH, immunofluorescence and PLA experiments, performed ChIP-Seq and ChIP-qPCR in HepG2 cells and mESC cells and did western blot and qRT-PCR in HepG2 cells and mESC cells. I performed and imaged FISH and immunofluorescence experiments, developed and applied automatic image analysis pipelines for the quantification of FISH, immunofluorescence and PLA experiments, performed ChIP-qPCR and western blot in MEF cells and FAIRE in HepG2 cells. Roberto Malinverni analyzed ChIP-Seq data and performed the association analysis. Justine

Renauld and Marc Thiry obtained and analyzed electron microscopy images. Vanesa Valero replicated and provided support in many of the experiments. Fabien Mongelard and Philippe Bouvet obtained and provided the MEF cells from control and *macroH2A1* -/- mice. Together with Julien Douet and Marcus Buschbeck, I contributed to the analysis and interpretation of all data sets.

7. Results II: MacroH2A can limit chromatin plasticity through two distinct mechanisms

7.1. MacroH2A proteins show different tissue expression patterns and conserved structural features

Despite the high degree of sequence similarity between macroH2A1.1, macroH2A1.2 and macroH2A2, several studies have demonstrated major functional differences between them. The analysis of the expression of macroH2A1.1, macroH2A1.2 and macroH2A2 in a panel of adult mouse tissues readily reveals a complex expression pattern with different relative levels of each protein in different tissues (**Figure 27A**). For example, macroH2A2 is highly expressed in the adult mouse brain (**Figure 27A**) in relation to macroH2A1.1 and macroH2A1.2. In the maturing mouse embryo, mRNA for macroH2A2 is strongly detected in the central nervous system (**Figure 27B**). Analysis of human samples shows fine-tuned expression differences between cell types in the same tissue. In human testis, macroH2A1.1 is distinctly detected in post-mitotic Sertoli cells, while macroH2A1.2 is also present in proliferating germ cells (**Figure 27C**). Similarly, in human lymph nodes macroH2A1.1 expression is low in highly proliferating cells that form the inner lymph node but macroH2A1.2 is detected across the tissue (**Figure 27C**). These findings strengthen previous studies that observed an inverse correlation of macroH2A1.1 expression with cell proliferation (Chen et al., 2014; Gamble et al., 2010; Sporn et al., 2009).

Results

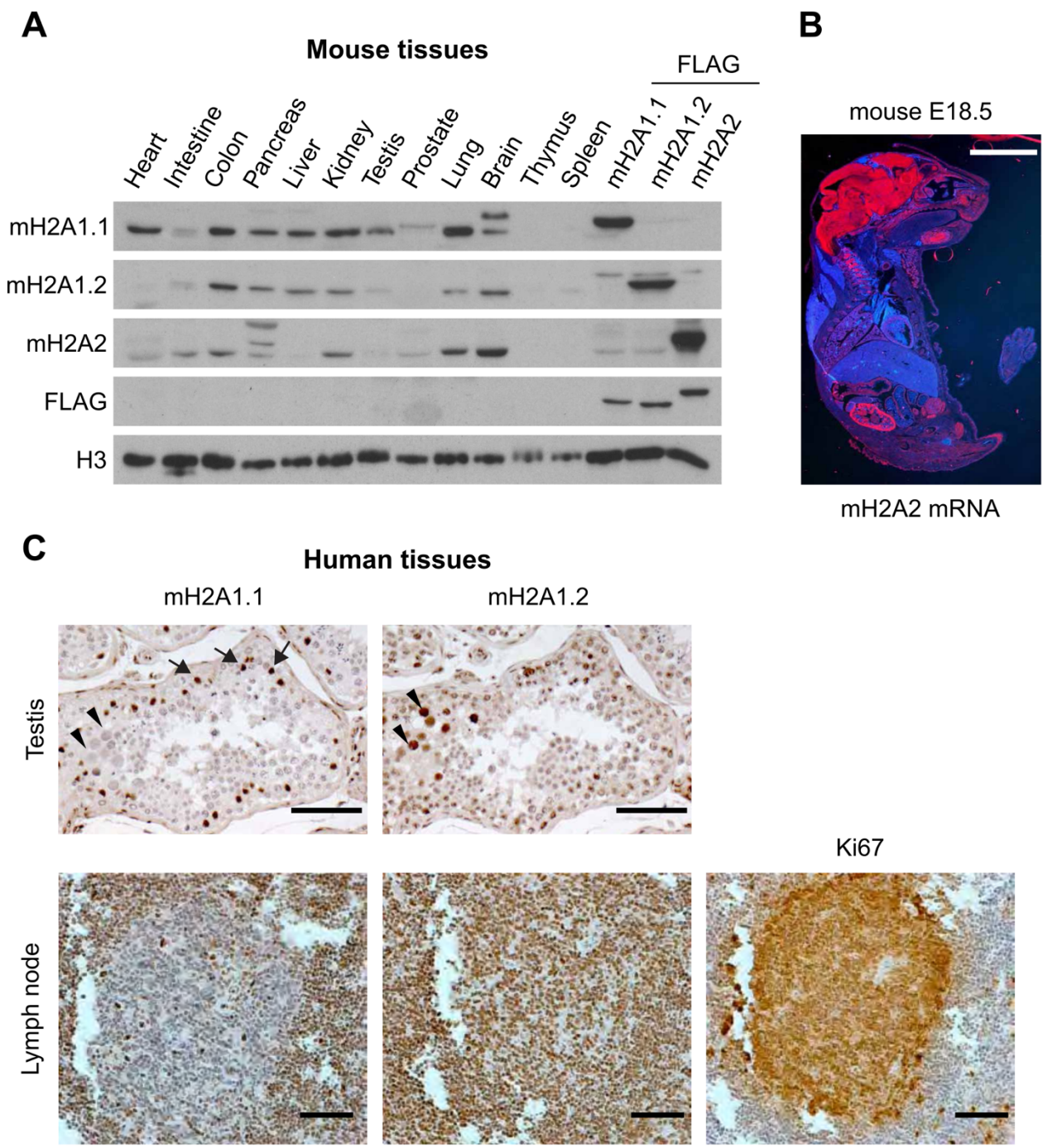


Figure 27. MacroH2A histone variants show a tissue- and cell type-specific expression pattern. (A) Western blot showing the expression of macroH2A1.1, macroH2A1.2 and macroH2A2 in different adult male mouse tissues. Samples of HepG2 cells expressing ectopic FLAG-tagged forms of the three macroH2A proteins are included as antibody specificity controls and to aid comparisons. H3 is used as a loading control. **(B)** In situ mRNA hybridization of mouse macroH2A2 on 20 µm cryosections of E18 mice reveals a high presence of macroH2A2 mRNA in the central nervous system, kidney and testis. Scale bar is 5 µm. **(C)** Immunohistochemical staining of macroH2A1.1 and macroH2A1.2 in paraffin-embedded human testis and lymph node sections. In testis, macroH2A1.1 is distinctively expressed in Sertoli cells (arrows) while macroH2A1.2 is also expressed in proliferative germ cells (arrow heads). In the lymph node, cells positive for the proliferation marker Ki67 express low levels of macroH2A1.1 while macroH2A1.2 is detected across the whole tissue section. Scale bars are 50 µm.

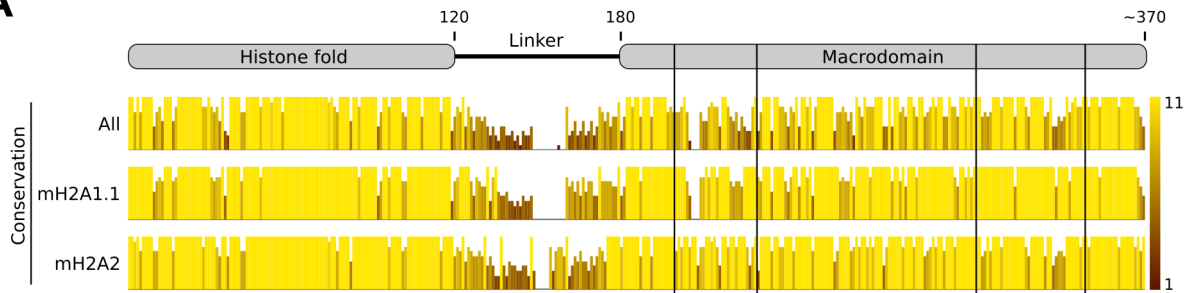
The three members of the macroH2A family of histone variants have a highly similar sequence and tertiary structure and are well conserved throughout evolution. The alignment of human (*H. sapiens*), zebrafish (*D. rerio*) and medaka (*O. latipes*) macroH2A1.1 and macroH2A2 amino acid sequences shows an overall high conservation of the histone fold and macrodomain, while the linker region between them is much less conserved (**Figure 28A**). One of the most notable features of the macrodomain is its pocket which, in the case of macroH2A1.1 proteins, is able to bind ADP-ribose. Despite the high conservation of the macrodomain sequence, some conserved divergences exist between macroH2A1.1 and macroH2A2 in the surface of this pocket. The amino acids that directly interact with ADP-ribose are highly conserved in macroH2A1.1 (**Figure 28B**). On the other hand, macroH2A2 proteins from the three analyzed species present a proline that distorts the phosphate-binding loop (Pro315), a three amino acid insertion in the adenine-binding site and a charged aspartate or glutamate at the position corresponding to glycine 224 in macroH2A1.1 (**Figure 28B**). The mutation of Gly 224 for glutamate in macroH2A1.1 is known to abolish ADP-ribose binding (Kustatscher et al., 2005).

Despite these differences and consistent with the ~65% amino acid sequence identity between macroH2A1.1 and macroH2A2, the 1.7 Å crystal structure of the macroH2A2 macrodomain shows a conserved globular macrodomain fold (**Figure 29A, B**). However, the amino acid changes in the macrodomain pocket when compared to macroH2A1.1 provoke major differences in the shape and chemical properties of this part of the protein. Specifically, the presence of a proline (Pro 315) instead of the glycine found in macroH2A1.1 (Gly 312) distorts the P-loop-like motif that coordinates the binding of the phosphates of the ADP-ribose in macroH2A1.1, which makes the structure of macroH2A2 incompatible with nucleotide binding (**Figure 29B, C**). Interestingly, in contrast to the splice isoform macroH2A1.2, the macrodomain of macroH2A2 has a conserved aromatic group (Phe354) that in macroH2A1.1 (Phe351) is involved in an aromatic stacking interaction with the nucleotide base of ADP-ribose (**Figure 29B, Figure 29C, D**) (Kustatscher et al., 2005).

Taken together, these conserved sequence divergences point at the inability of macroH2A2 to bind ADP-ribose, a highly conserved function of the macroH2A1.1 macrodomain. However, some structural features of macroH2A conserved in vertebrate evolution suggest that it could potentially bind a nucleobase or related metabolite molecule.

Results

A



B

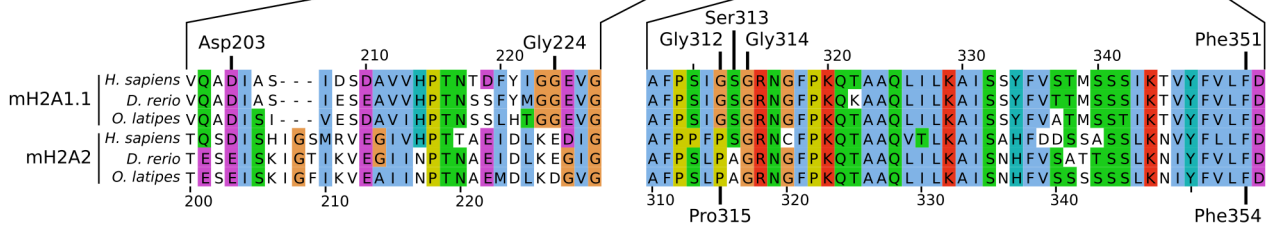


Figure 28. MacroH2A1.1 and macroH2A2 show conserved divergences in vertebrate evolution. (A)

Conservation scores of performing multiple sequence alignment of macroH2A1.1 and macroH2A2 protein sequences of human (*H. sapiens*), zebrafish (*D. rerio*) and medaka (*O. latipes*) all together or separately with macroH2A1.1 and macroH2A2 sequences. The conservation score reflects the conservation of physico-chemical properties of the amino acids in each position of the alignment in a scale that ranges from 0 to 11 and is represented in both the height and color of the bar. The schematic at the top displays the different domains of macroH2A histones. The alignment was generated with the online alignment tool PRALINE [REF] using the protein sequences with the following accession codes: *H. sapiens* macroH2A1.1 (NP_613075.1), *D. rerio* macroH2A1 (NP_001035451.1), *O. latipes* macroH2A1.1 (XP_011481424.1), *H. sapiens* macroH2A (NP_061119.1), *D. rerio* macroH2A2 (NP_001020673.1) and *O. latipes* macroH2A2 (XP_004076965.2). The alignment and conservation track were edited with Jalview (Waterhouse et al., 2009). **(B)** Protein sequence alignment of the amino acids that form the pocket of the macrodomain of macroH2A1.1 and macroH2A2 from human (*H. sapiens*), zebrafish (*D. rerio*) and medaka (*O. latipes*). Critical amino acids for ADP-ribose binding conserved in macroH2A1.1 are indicated at the top, as well as the amino acids in macroH2A2 cited in the main text at the bottom. The top and bottom amino acid sequences refer to human macroH2A1.1 and macroH2A2 sequences, respectively. The alignment was generated as in (A) and colored using the ClustalX color scheme available in Jalview. In summary, amino acids are colored according to their physico-chemical properties only if that position in the alignment fulfils a conservation criteria for that amino acids type: blue - hydrophobic, red - positive charge, magenta - negative charge, green - polar, pink - cysteine, orange - glycine, yellow, proline, cyan - aromatic, white - no criteria met.

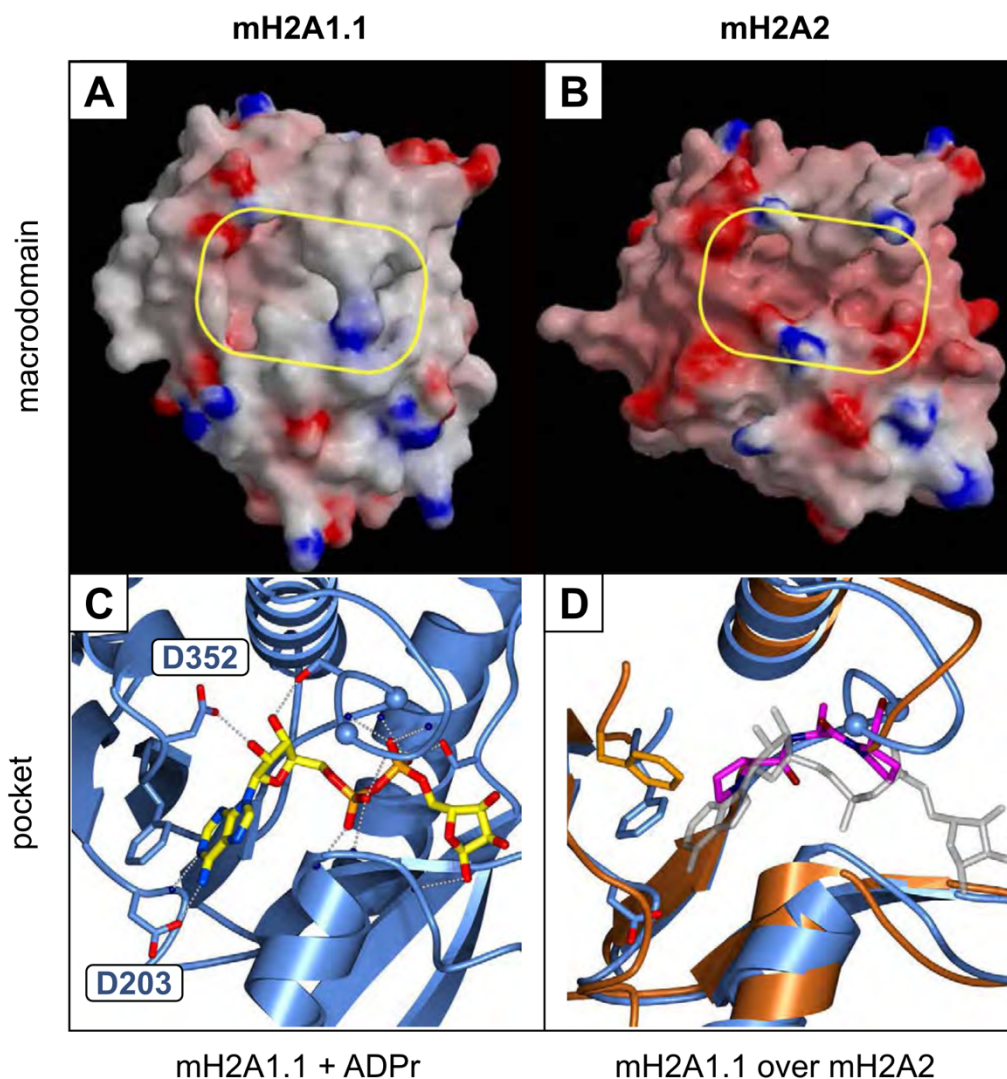


Figure 29. MacroH2A proteins show specific structural differences in their macrodomains. (A, B) Surface representations of the macrodomain structures for human macroH2A1.1 (PDB ID 1ZR3) and macroH2A2 (PDB ID 6FY5). The coloring represents the electrostatic potential (red - positive, blue - negative, grey - neutral). The yellow frame indicates the major surface pocket which in the case of macroH2A1.1 can accommodate ADP-ribose. **(C)** Close-up view of the nucleotide binding region of the ADP-ribose molecule (yellow) bound by macroH2A1.1 (blue) at its major surface pocket. The side chains of two important aspartate residues are marked. Glycines 312 and 314 in the nucleotide binding loop are depicted as balls. Dashed lines indicate polar interactions. **(D)** Overlay of the macroH2A1.1 pocket region depicted in (C) in blue with the corresponding region in macroH2A2 in orange. The ADP-ribose ligand shaded in light grey is included to aid comparison and orientation. Important structural differences in the distort the nucleotide binding loop in macroH2A2 are highlighted in magenta.

7.2. MacroH2A1.1, but not macroH2A2, binds ADP-ribose and inhibits PARP1

We used isothermal titration calorimetry (ITC) to measure the capacity of macroH2A1.1 and macroH2A2 to bind ADP-ribose. Consistent with the structural analysis, the macrodomain of macroH2A1.1 bound monomeric ADP-ribose *in vitro* while the macrodomain of macroH2A2 was unable to do so (**Figure 30A**). The capacity of macroH2A1.1 to bind ADP-ribose is conserved between the human protein and the orthologs in zebrafish and medaka (**Figure 30B**). This functional conservation of the macrodomain module of macroH2A1.1 coincides with the high degree of sequence conservation, particularly in its major surface pocket (**Figure 28**). Previous structural analysis of the macroH2A1.1 macrodomain in complex with monomeric ADP-ribose concluded that this binding event could allow macroH2A1.1 to “cap” oligo-ADP-ribose chains (Timinszky et al., 2009). This mechanism is proposed to explain the capacity of macroH2A1.1 to bind and inhibit auto-modified PARP1 (Ouararhni et al., 2006; Posavec-Marjanović et al., 2017). Some contradictory results exist showing that other macrodomain modules, and in particular the macrodomain of macroH2A2, can inhibit PARP1 activity (Nusinow et al., 2007b). To investigate this issue, we directly compared the capacity of the macrodomains of macroH2A1.1 and macroH2A2 to inhibit PARP1 enzymatic activity in an *in vitro* assay (**Figure 30C**). Using this approach, we found that increasing amounts of the macrodomain macroH2A1.1 progressively decreased the activity of PARP1 *in vitro* (**Figure 30D**). In contrast, PARP1 enzymatic activity was not affected by the addition of similar amounts of the macroH2A2 macrodomain or the macroH2A1.1 G224E mutant unable to bind ADP-ribose (**Figure 30D**). This demonstrates that the repressive capacity of macroH2A over PARP1 is restricted to macroH2A1.1 and is completely dependent on the ability to bind ADP-ribose through its macrodomain.

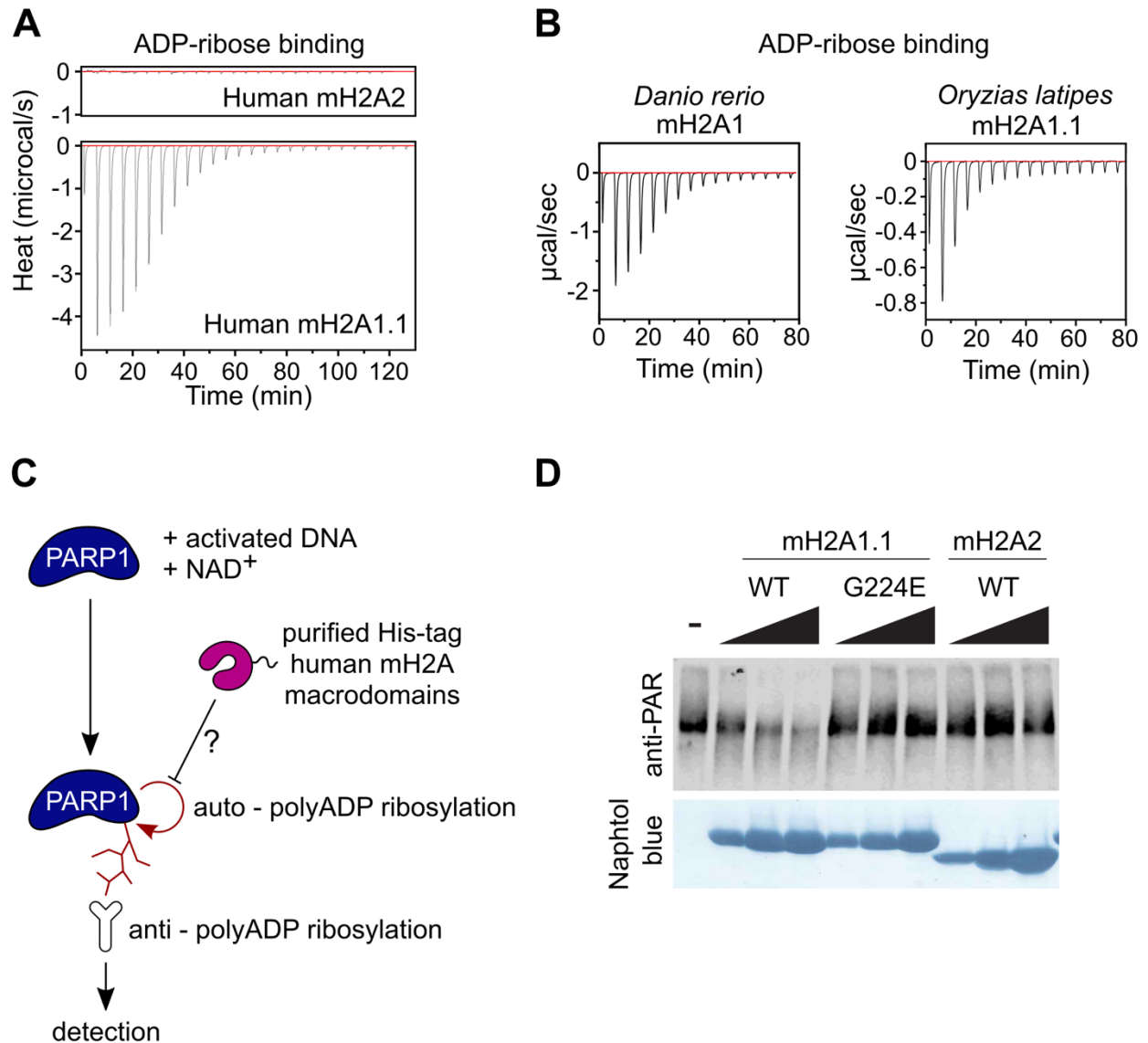


Figure 30. The conserved binding of ADP-ribose to macroH2A1.1 is essential of its unique repressive effect on PARP1. (A) Isothermal calorimetry (ITC) assays using ADP-ribose and purified macrodomains of human macroH2A1.1 and macroH2A2. ITC is based on the measure of heat absorption by binding events upon gradual titration of the ligand over time. In the case of macroH2A2 macrodomain, the flat profile indicates that no binding is occurring. On the other hand, addition of ADP-ribose to the sample containing macroH2A1.1 macrodomains results in heat absorption that diminishes as the binding sites are occupied and eventually become saturated. (B) ITC assays using ADP-ribose and purified macroH2A1.1 macrodomains of zebrafish (*Danio rerio*) and medaka (*Oryzias latipes*) shows conserved binding of ADP-ribose. Note that in zebrafish only one macroH2A1 splice variant is present (ortholog to macroH2A1.1) and is referred to as simply macroH2A1. (C) Schematic of the in vitro assay used to evaluate the repressive capacity of macroH2A on PARP1. PARP1 is put in the sample on the presence of activated DNA and NAD^+ , which turns it enzymatically active, and purified human macrodomains. The first step in PARP1 activation involves its auto modification by ADP-ribosylation and subsequent synthesis of poly(ADP-ribose) chains, which can be detected with a specific antibody for this PTM in a western blot. Thus, the signal strength in the western blot is effectively a measure of PARP1 enzymatic activity in the sample. (D) In vitro PARP1 activity assay in the presence of increasing concentrations (10, 25 and $50\mu\text{M}$) of purified macrodomains of human macroH2A1.1, macroH2A1.1 G224E mutant and macroH2A2. The naphthol staining shows the increasing amounts of pf purified macrodomain in each reaction. A representative blot of three independent experiments is shown.

7.3. MacroH2A histones suppress chromatin dynamics upon DNA damage through two distinct mechanisms

The ability of macroH2A1.1 to repress PARP1 activity *in vitro* raises the question as to whether this histone variant can suppress chromatin dynamic processes involving the PARP1 enzyme. As a model to study this effect, we used the rapid decondensation of chromatin that is observed at and near DNA damage sites *in vivo*. Previous work has shown that this structural rearrangement of chromatin depends on PARP1 activity (Luijsterburg et al., 2016; Sellou et al., 2016; Strickfaden et al., 2016). To be able to study and measure this process in live cells, we expressed a photo-activable GFP (PAGFP) fused to histone H2B (Kruhlak et al., 2006) in human U2OS osteosarcoma cells that had been sensitized to DNA damage by pre-incubation with the DNA-intercalating Hoechst dye. Microirradiation in the cell nucleus with a 405nm laser simultaneously induces DNA damage and photo-activation of the PAGFP-H2B (**Figure 31A**), allowing the measurement of the extent of chromatin expansion using live cell imaging (**Figure 31B**). Importantly, PAGFP was shown to remain bound to chromatin during this process (Kruhlak et al., 2006), which makes the fluorescent signal of PAGFP an effective marker to allow the tracking of chromatin at the site of irradiation. The rapid expansion that occurs at the DNA damage site is completely suppressed by pre-treating the cells with the PARP1 inhibitor Olaparib, validating the dependence of the process on PARP1 enzymatic activity (**Figure 31C**).

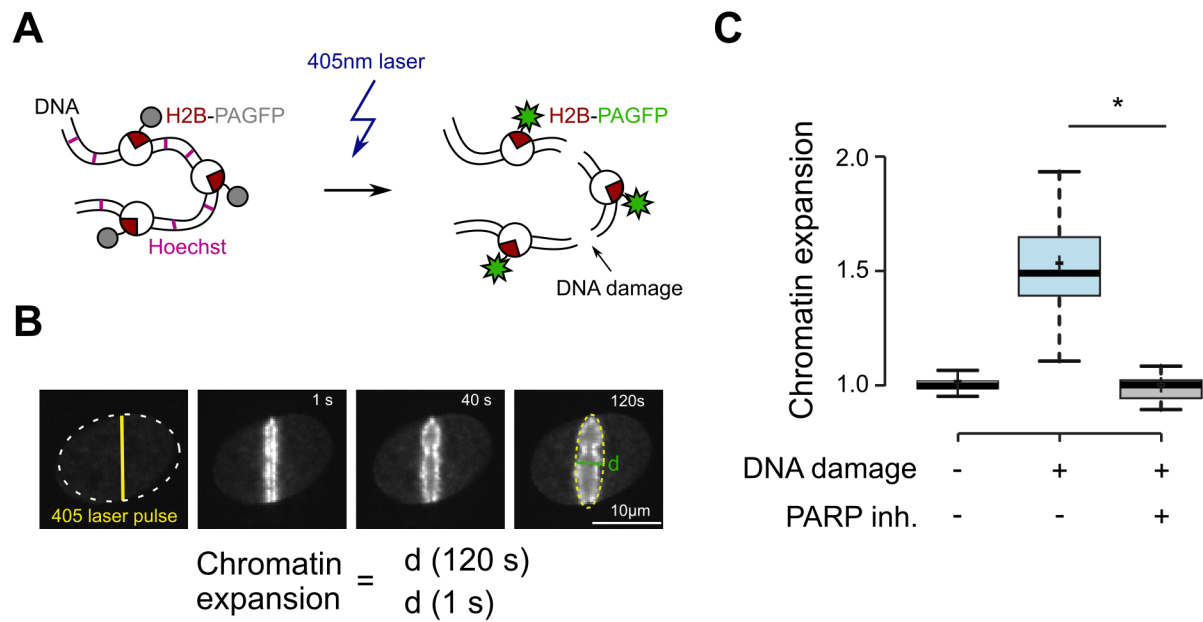


Figure 31. Chromatin rearrangement at and near DNA damage sites depends on the enzymatic activity of PARP1. (A) Schematic of the system used to induce simultaneous DNA damage and chromatin-bound PAGFP activation in living cells with the combination of a DNA intercalating Hoechst dye and a 405nm laser pulse. (B) Example fluorescence images of the model used to study chromatin dynamics upon DNA damage. The DNA intercalating agent Hoechst and a 405nm laser microirradiation are used to induce DNA damage. The PAGFP-H2B proteins at the damaged site are photo-activated by the laser, which allows the tracking over time of chromatin expansion. In particular, we have measured the width of the microirradiated chromatin (d) as the minor axis of an ellipsoid fitted to the area of the fluorescent signal of PAGFP-H2B. The measurements at 120s are normalized by the 1s timepoint to achieve the final "chromatin expansion" measure. (C) Chromatin expansion in the absence and presence of intercalating Hoechst dye (DNA damage - and +) and with or without pre-treatment for 1h with 1µM of the PARP1 inhibitor Olaparib (PARP inh. - and +). Single cell measurements of three biological replicates are plotted with $n > 30$ each. A two-tailed Student's T-test was used, * $p < 0.05$.

Results

To compare the effect of the different macroH2A histones on this chromatin rearrangement process, we transiently overexpressed mCherry-tagged human histone constructs in the U2OS cells stably expressing PAGFP-H2B (**Figure 32A**). We used fluorescence recovery after photobleaching (FRAP) assays to test the incorporation of these ectopic constructs into the chromatin (Bönisch et al., 2012). Both the H2A and macroH2A fluorescent constructs have very low mobility in living cells, indicating that they are incorporated into chromatin (**Figure 32B**). We then performed the laser-induced DNA damage assay described before (**Figure 31A, B**) to monitor the effect of overexpressing each mCherry-tagged histone in the extent of chromatin expansion. The mCherry fluorescence allowed to select cells with similar levels of expression of the different constructs for this assay. We found that overexpressing canonical H2A had no effect while macroH2A1.1 had the strongest effect in suppressing chromatin expansion (**Figure 33A**), which correlates with its ability to inhibit PARP1. The G224E mutation in macroH2A1.1, which abolishes ADP-ribose binding and PARP1 inhibition, reduced the inhibitory effect of macroH2A1.1 but did still produce a lower chromatin expansion when compared with the H2A control (**Figure 33A**). Similarly, the splice isoform macroH2A1.2 and macroH2A2 also significantly reduced chromatin expansion when compared to H2A but to a lesser extent than wild-type macroH2A1.1 (**Figure 33A**). To rule out interference of the fluorescent tag with chromatin dynamics, we have confirmed that the inducible expression of macroH2A1.1 in HeLa cells without any fluorescent tag produces similar results in terms of reducing chromatin expansion (**Figure 33B**).

In summary, we found that macroH2A1.1 is the macroH2A histone with the strongest effect in suppressing DNA-damage induced chromatin expansion, which reflects its capacity to inhibit the activity of PARP1. However, macroH2A1.2 and macroH2A2, as well as the macroH2A1.1 G224E mutant, all show a smaller but significant capacity to reduce chromatin expansion despite their inability to interfere with PARP1 activity. This indicates the existence of at least a second mechanism by which macroH2A can suppress chromatin dynamics.

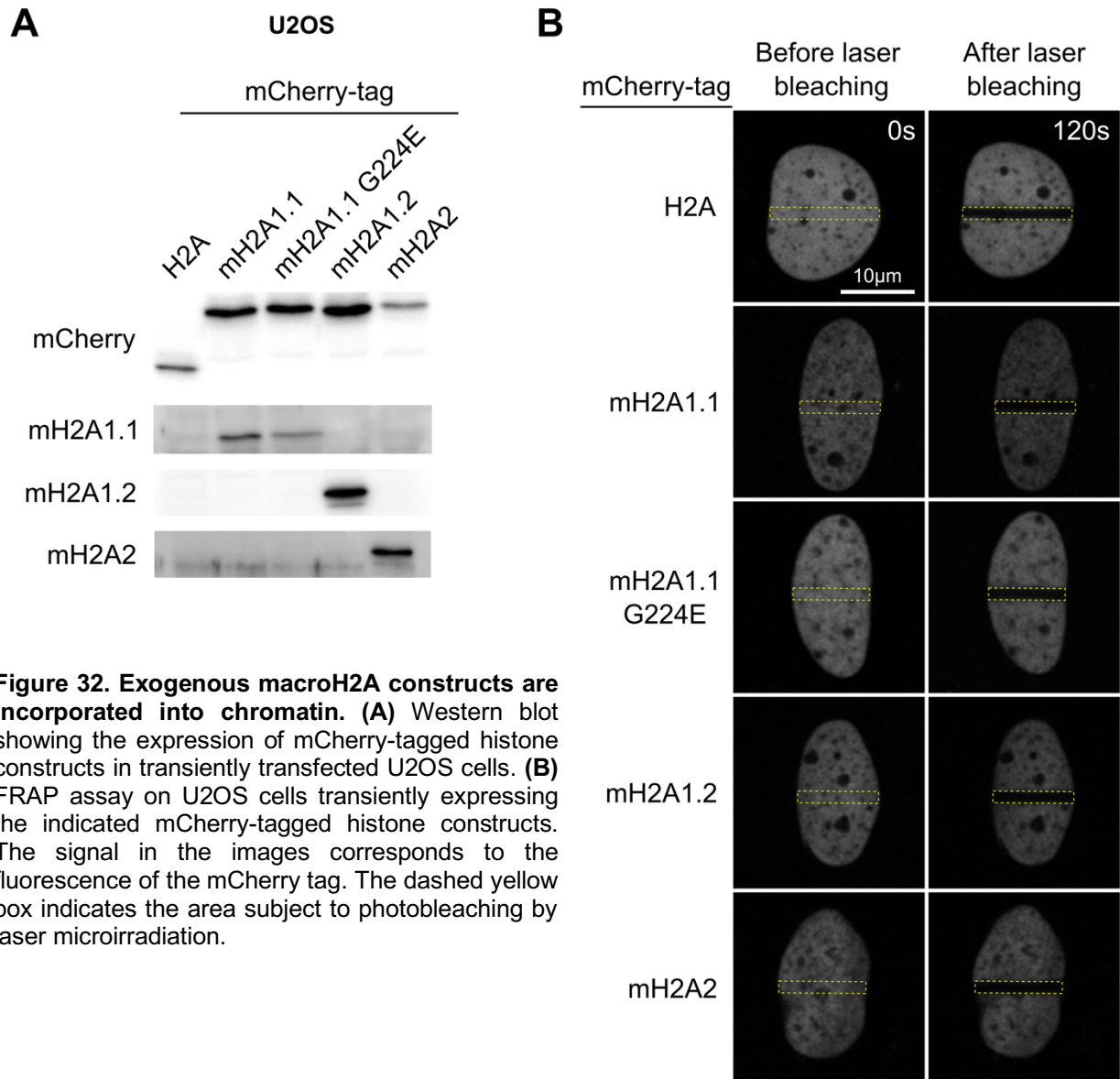


Figure 32. Exogenous macroH2A constructs are incorporated into chromatin. (A) Western blot showing the expression of mCherry-tagged histone constructs in transiently transfected U2OS cells. **(B)** FRAP assay on U2OS cells transiently expressing the indicated mCherry-tagged histone constructs. The signal in the images corresponds to the fluorescence of the mCherry tag. The dashed yellow box indicates the area subject to photobleaching by laser microirradiation.

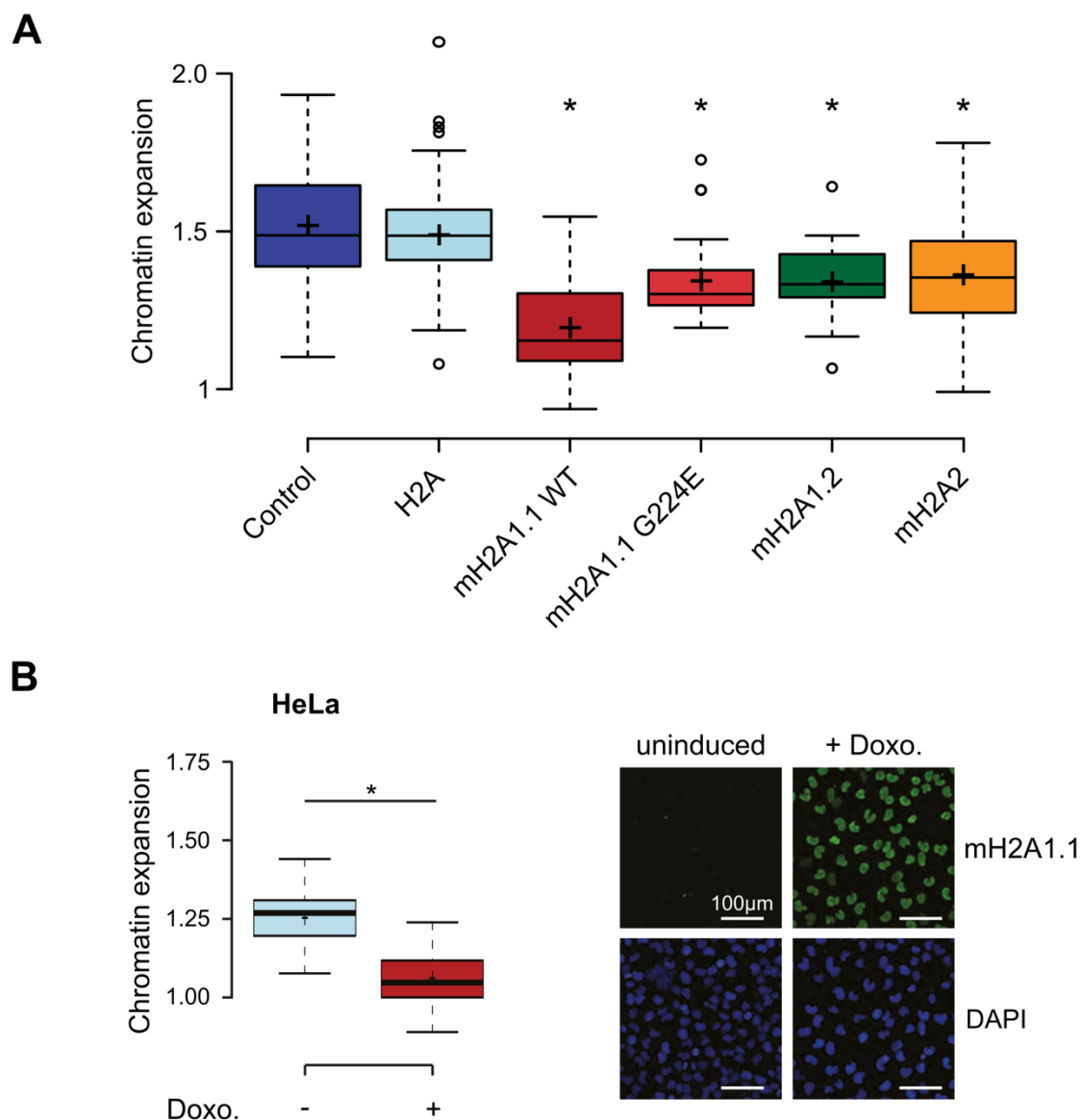


Figure 33. MacroH2A histones repress chromatin expansion at DNA damage sites, macroH2A1.1 having the strongest effect. (A) Quantified chromatin expansion (as described in Fig 15A, B) in living U2OS cells not transfected (control) or transiently expressing each indicated mCherry-tagged histone. Data is plotted as single cell measurements of three biological replicates each with $n > 30$. A two-tailed Student's T-test was used to compare each group with cells expressing mCherry-H2A, * $p < 0.05$. **(B)** Chromatin expansion measured in HeLa cells containing a doxocycline (doxo.) inducible macroH2A1.1 transgene. The chromatin expansion assay (left panel) was performed 48h after induction and the expression of ectopic macroH2A1.1 was monitored with an anti-macroH2A1.1 immunofluorescence (right panel). Data representation and statistics as in (A).

7.4. The linker region of macroH2A limits chromatin expansion upon DNA damage in living cells

To investigate the nature of this second mechanism, we interrogated which domains of macroH2A are involved in the suppression of chromatin expansion. MacroH2A histone variants have a tripartite structure composed of: an N-terminal histone fold, an unstructured linker and a C-terminal macrodomain (**Figure 7**). We generated truncation mutants with deletions of either the macrodomain or the macrodomain and the linker of histones macroH2A1.1 and macroH2A2 (**Figure 34A**) and transiently expressed them in U2OS cells that were subject to the same chromatin expansion assay described before. We found that both macroH2A1.1 and macroH2A were able to reduce chromatin expansion after the deletion of the entire macrodomain (**Figure 34B**). However, this capacity was abolished by eliminating the unstructured linker domain of both proteins (**Figure 34B**). Notably, the strongest inhibitory effect in chromatin expansion in this experiment was still achieved by overexpressing the full-length macroH2A1.1 protein, highlighting the importance of the macrodomain module of this protein. Importantly, the linker region of macroH2A1.1 was sufficient to inhibit chromatin expansion when fused to the C-terminus of the canonical H2A histone (**Figure 34C**). Additionally, an even stronger suppressive effect could be achieved when fusing the linker and the macrodomain of macroH2A1.1 to canonical H2A (**Figure 34C**).

These experiments indicate that the basic unstructured linker region of all macroH2A histones is sufficient to reduce chromatin expansion upon acute DNA damage.

Results

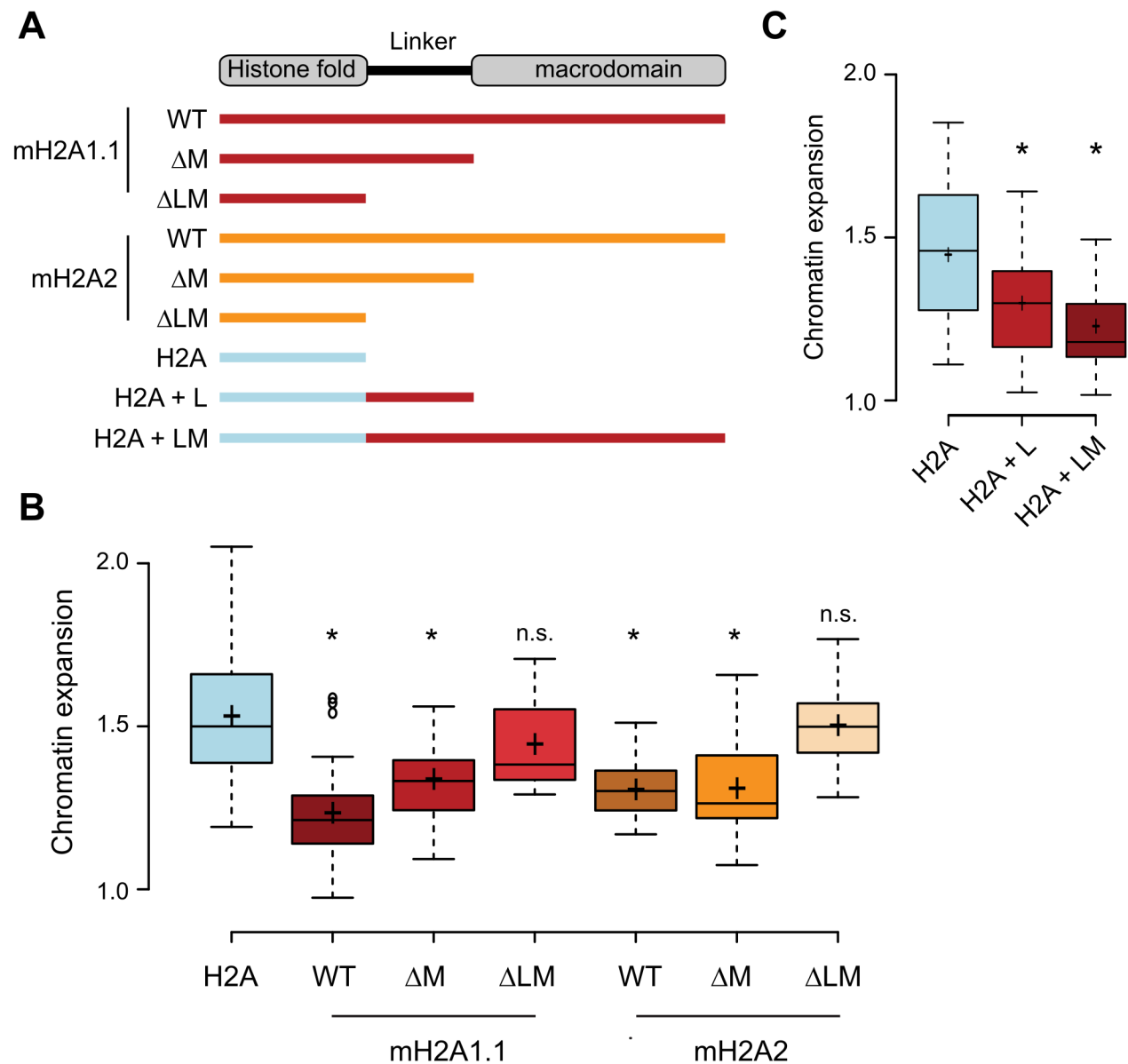


Figure 34. MacroH2A histones repress chromatin expansion via their basic linker and the macrodomain capable of ADP-ribose binding. (A) Schematic of the macroH2A and H2A constructs used in the study: full-length wild type (WT), truncation constructs with a deleted macrodomain (ΔM), truncation constructs with deleted macrodomain and linker (ΔLM) or the H2A fold fused to the linker of macroH2A1.1 (H2A + L) or the H2A fold fused to both the linker and macrodomain of macroH2A1.1 (H2A + LM). **(B, C)** Quantified chromatin expansion in cells transfected with the indicated histone constructs. Boxplots represent single cell measurements of three biological replicates with $n > 30$ each. A two-tailed Student's t-test was used to compare all groups with the H2A control, * $p < 0.05$, n.s not significant).

7.5. The macroH2A linker region is essential and sufficient for stabilizing heterochromatin architecture

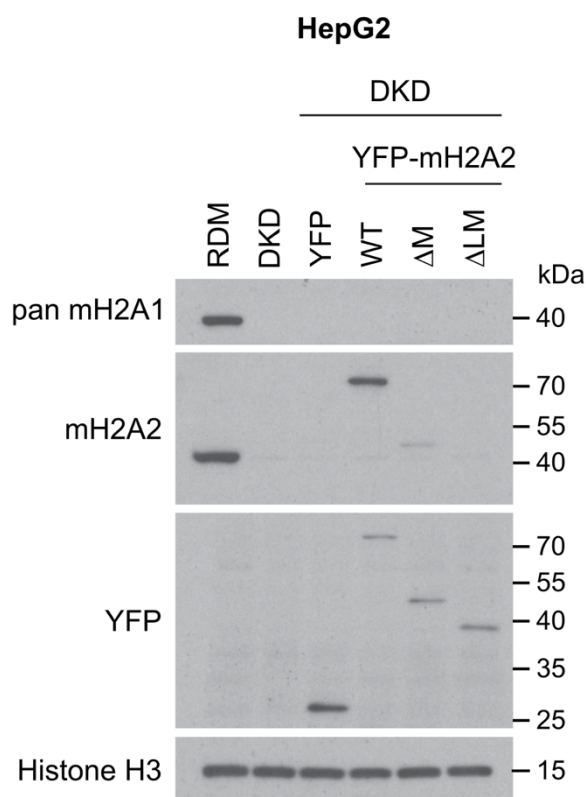


Figure 35. Exogenous expression of YFP-tagged macroH2A2 constructs in HepG2 cells. Western Blot showing the expression of macroH2A proteins in control HepG2 cells (RDM), knockdown for macroH2A1 and macroH2A2 (DKD) and DKD with ectopic expression of the indicated YFP-tagged constructs. The deletion mutants annotation follows the same as in Fig 18A. H3 is included as a loading control. The pan mH2A1 antibody detects both the macroH2A1.1 and macroH2A1.2 isoforms.

We have previously shown that loss of all macroH2A histones leads to major alterations in nuclear organization and heterochromatin architecture (6, Douet et al. 2017). This included an increase in nuclear size and dispersion of heterochromatic elements in HepG2 cells. We wished to test whether the linker domain of macroH2A contributes to these functions outside of the context of the DNA-damage response process. For this, we stably integrated YFP-tagged constructs of wild-type macroH2A2 and its deletion mutants in HepG2 cells that were depleted of the endogenous expression of macroH2A histones (**Figure 35**). We used confocal imaging to measure nuclear size by analyzing the DAPI staining and assessed heterochromatin architecture by quantifying the number of H3K9me3 foci detected in the nucleus by automatic segmentation (**Figure 36A-C**). We observed a bigger nuclear size, as well as an increase in the number of

H3K9me3 foci in the nucleus, in macroH2A-deficient DKD cells (**Figure 36A, B**). The expression of YFP-tagged full-length macroH2A rescued both nuclear size (**Figure 36A, B**) and heterochromatin structure (**Figure 36A, C**), while the expression of the fluorescence tag YFP alone did not have any effect. Next, we tested whether a truncated form of macroH2A2 lacking the macrodomain but still retaining the linker region was able to rescue this two phenotypes. We found that this truncated form was still capable of rescuing the function by reducing the nuclear size and number of H3K9me3 foci (**Figure 36A-C**). On the other hand, the truncation mutant lacking both the macrodomain and the linker, effectively containing only the histone fold of macroH2A2, was unable to rescue the observed phenotypes (**Figure 36A-C**).

Results

These results show that macroH2A proteins are involved in the regulation of chromatin structure in the context of constitutive H3K9me3-marked heterochromatin. In particular, the linker region of macroH2A2, but not its macrodomain, is necessary and sufficient to maintain nuclear size and heterochromatic organization.

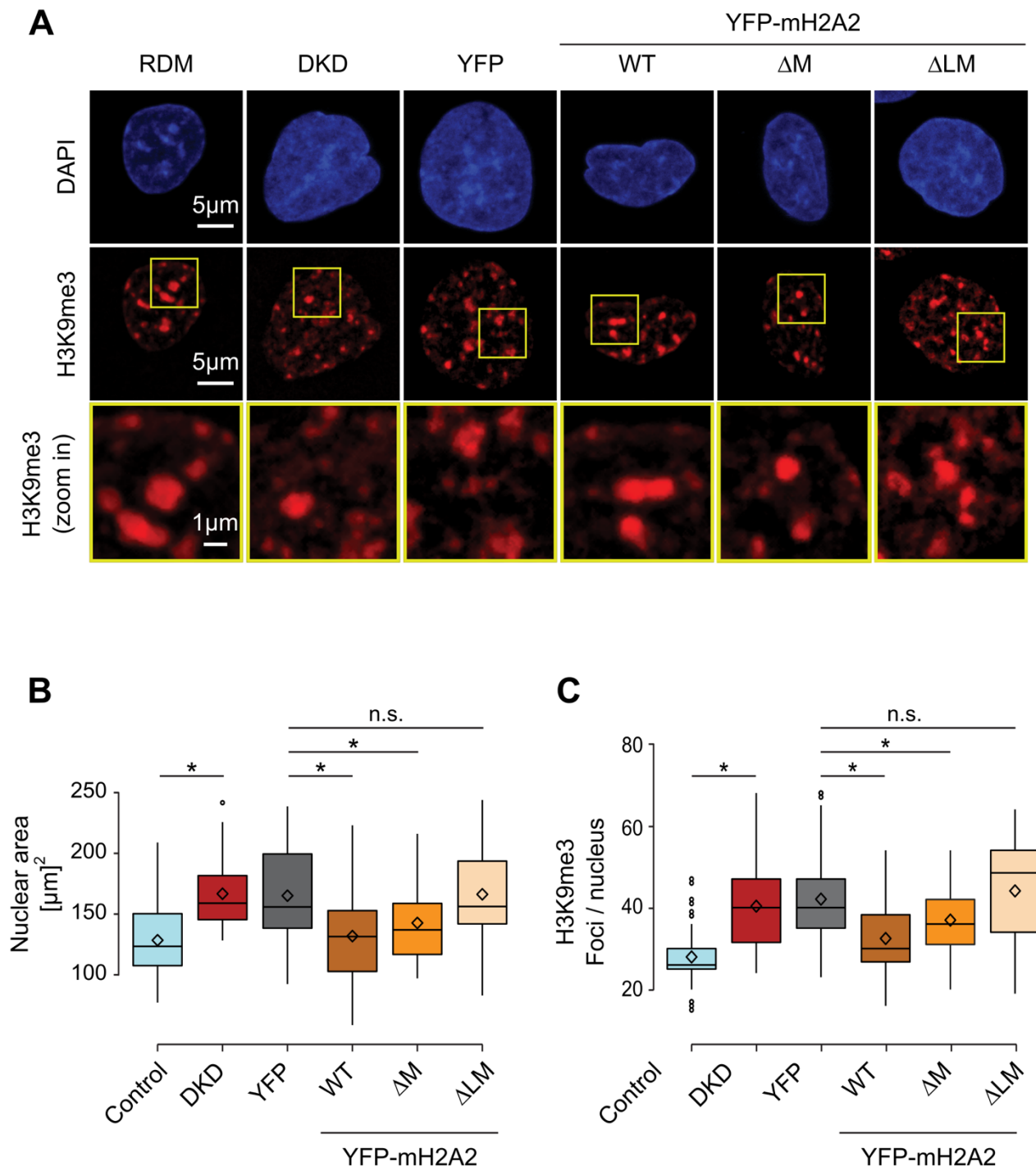


Figure 36. The linker domain of macroH2A2 is able to maintain heterochromatic structures. (A) H3K9me3 immunofluorescence (red) in the indicated HepG2 cell lines (Fig 19). The areas indicated with a yellow box are zoomed-in for a detailed observation of the H3K9me3 signal. Nuclei are counterstained with DAPI (blue). All images are maximum-intensity Z-stacks projections of confocal stacks. **(B, C)** Quantification of the nuclear area and number of H3K9me3 foci per nucleus, respectively. Data is plotted as single cell measurements from two biological replicates with $n > 50$. A two-tailed Student's T-test was used to make the indicated comparisons, * $p < 0.05$.

7.6. Contributions

This project was conceived and supervised by Andreas G. Ladurner, Marcus Buschbeck, Marek Kozlowski and myself. I generated stable HepG2 cell lines expressing exogenous macroH2A2 constructs, performed and imaged immunofluorescence experiments in HepG2, developed and applied automatic image analysis pipelines for the quantification of immunofluorescence experiments in HepG2 and performed the conservation analysis of the macrodomain. Marek Kozlowski, Imke Mandemaker and Charlotte Blessing performed and analyzed live cell imaging assays and western blot in U2OS cells. Michael Hothorn conducted the crystallographic experiments. Iva Guberovic performed the *in vitro* PARP1 activity assay. Arturo Gutiérrez-Triana performed the ITC assays. Judith Sporn conducted mRNA expression analysis in mouse and immunohistochemistry in human tissues.

8. Results III: MacroH2A represses the expression of DKK1 and attenuates its transcriptional activation by TNF α

8.1. HepG2 cells lacking macroH2A show an increased colony formation and migration capacity, but are unaffected in cell growth

The results presented in the two previous chapters characterized the role of macroH2A histone variants in chromatin structure, particularly in maintaining heterochromatin organization and modulating chromatin rearrangement at DNA-damage sites. Here, we address the question of whether macroH2A proteins are involved in gene regulation and if this is relevant for the pathogenesis of cancer cells.

To study this, we used our cellular model of HepG2 cell lines: control (RDM) or knockdown for macroH2A1 and macroH2A2 (DKD). Because HepG2 is a cancer cell line derived from hepatoblastoma, we first evaluated if the loss of macroH2A impacted some malignant properties associated with cancer cells using cell-cultured based assays. While the loss of macroH2A did not impact cellular proliferation in normal cell culture conditions (**Figure 37A**), DKD cells showed an increased capacity in colony formation when seeded at low cell density (**Figure 37B**). In addition, the DKD cells lacking macroH2A were able to migrate faster in wound healing assays (**Figure 37C, D**).

8.2. Tumors derived from macroH2A-lacking cells do not have increased growth or higher levels of major malignancy markers

Increased clonal growth capacity and migration are both malignancy-associated characteristics and indicate that DKD cells could behave as more aggressive cancer cells. To test this hypothesis, we injected HepG2 RDM and DKD cells in immunosuppressed mice to monitor their growth and resected the resulting tumors to subject them to histological and molecular analysis (**Figure 38A**). The samples corresponding to tumors derived from RDM and DKD HepG2 cell lines will be labelled with an X prefix (for xenograft): X.RDM and X.DKD (**Figure 38A**). We did not observe any difference in tumor growth between RDM and DKD cells when monitored over 21 days (**Figure 38B**). Importantly, we validated that both the expression of macroH2A1 and macroH2A2 genes was suppressed after the cells were grown in the mice (**Figure 38C**). Next, we performed an analysis of the histological characteristics of the generated tumors. The hematoxylin-eosin staining did not reveal differences between the X.RDM and X, both conditions

Results

showed an undifferentiated and pleomorphic tumor pathology with high mitotic activity (**Figure 38C**). Similarly, immunohistochemistry staining showed equal levels of the proliferation marker Ki67 and the liver tumor marker alpha-fetoprotein (AFP) (**Figure 38C**).

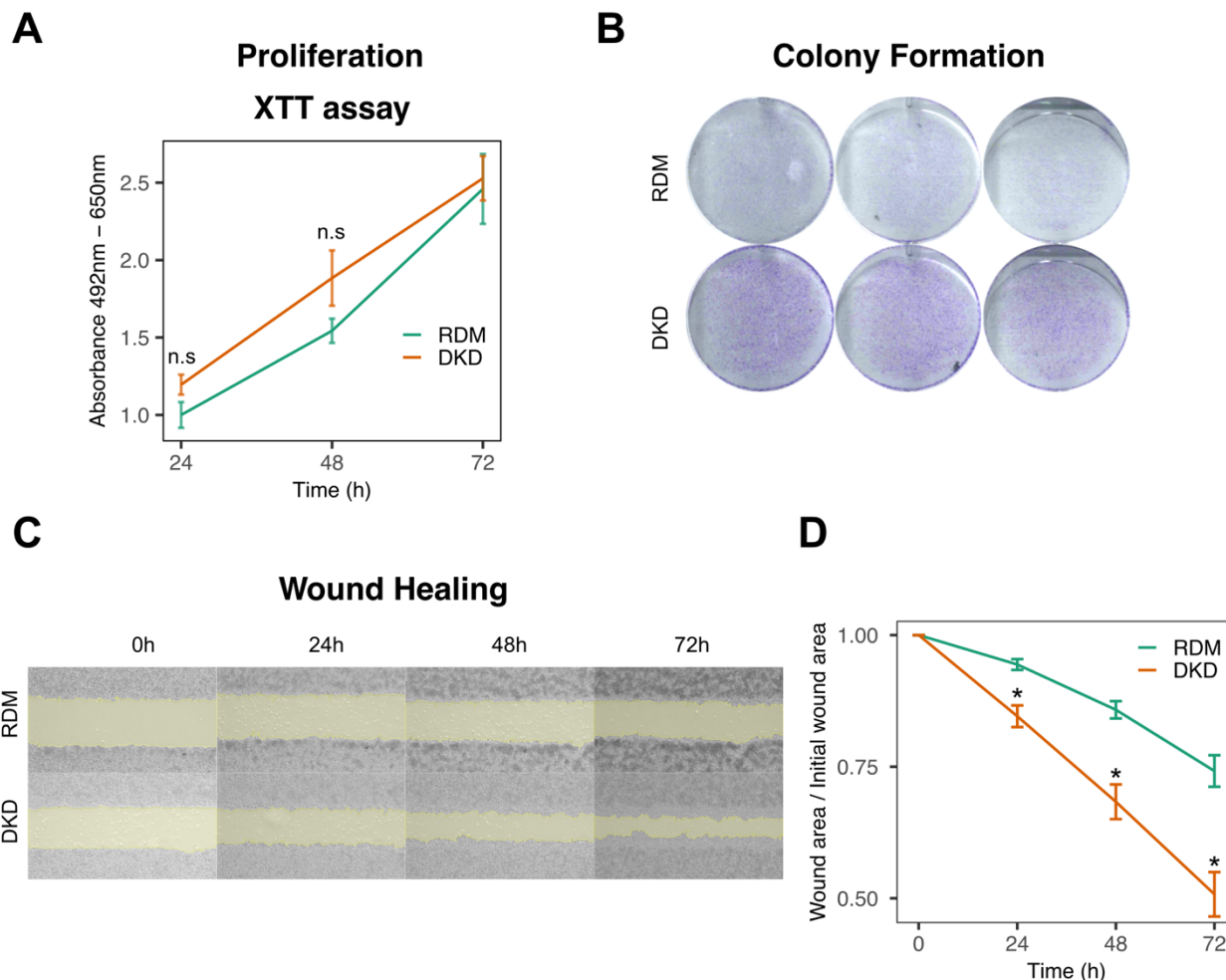


Figure 37. Cells lacking macroH2A have an increased colony formation and migration capacity. (A) Cell proliferation measured by XTT assay in HepG2 RDM and DKD cells. The XTT assay is a colorimetric assay based on the cleavage of tetrazolium salts exclusively in viable cells, thus obtaining a colorimetric signal proportional to the number of viable cells. A two-tailed Student's t-test was used to compare RDM and DKD at each timepoint, n.s not significant (n = 3). **(B)** Colony formation assay where the cells were seeded at low density, grown for one week and stained with crystal violet. Each column is a representative image from one biological replicate. **(C)** Representative images of a wound healing experiment as performed in HepG2 cells at the indicated timepoints. The wound area as measured by computational analysis is shaded in yellow. **(D)** Quantification of the migration capacity of HepG2 cells in wound healing assays measured as the wound area at a particular timepoint divided by the wound area at the initial (0h) timepoint. A two-tailed Student's t-test was used to compare RDM and DKD at each timepoint, * p < 0.05 (n = 3).

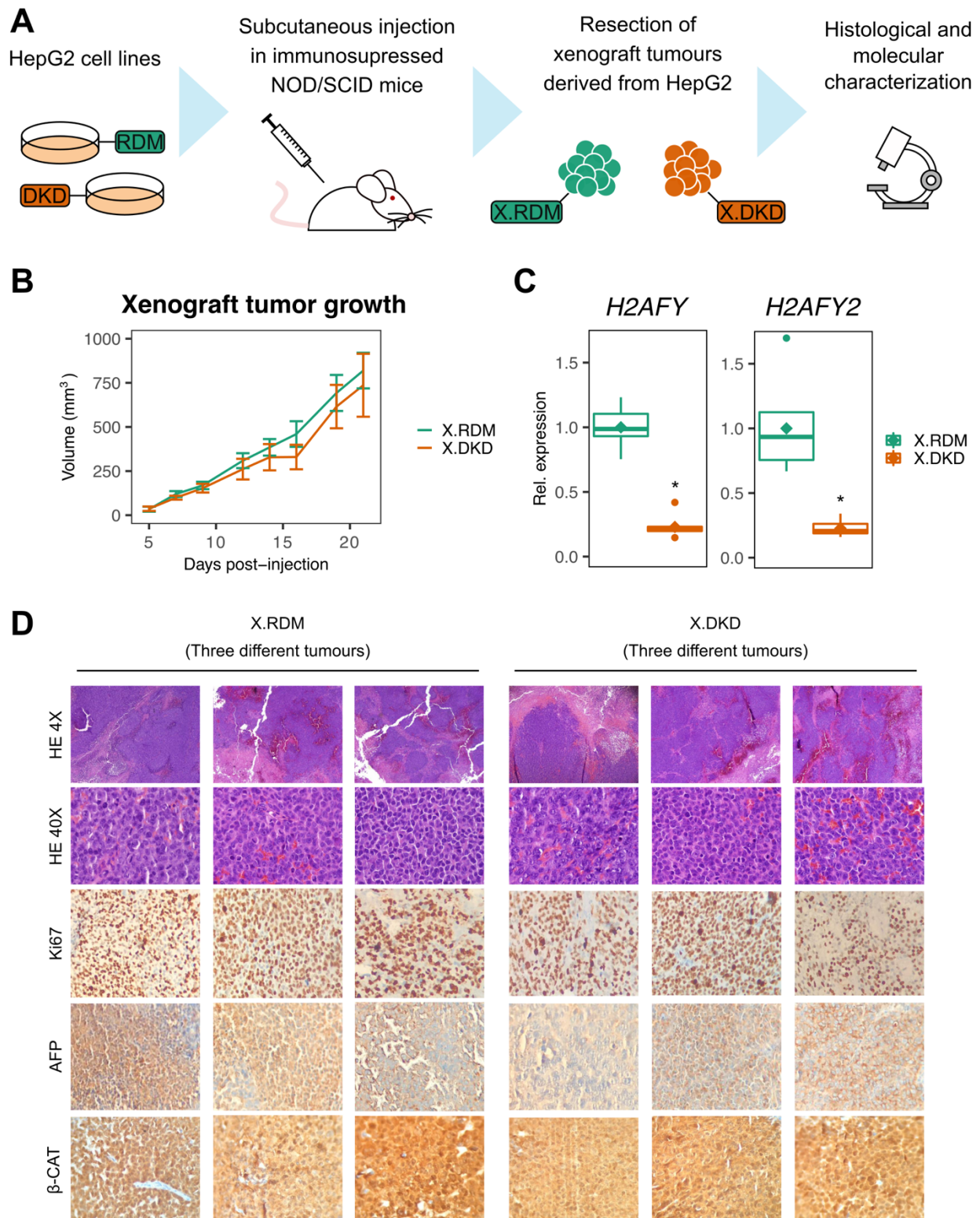


Figure 38. Loss of macroH2A in HepG2 cells does not result in increased tumor growth or alteration in major pathological markers. (A) Schematic of the experiment performed to evaluate the capacity of HepG2 cells to grow as tumors. HepG2 cell lines control (RDM) or lacking both macroH2A1 and macroH2A2 where subcutaneously injected in immunosuppressed NOD/SCID mice. Their growth was monitored in vivo and the resulting tumors were resected and subject to histological and molecular characterization. (B) Growth of xenograft tumors measured as tumor volume in vivo over 21 days after injection of the cells. Data is plotted as the mean of 10 different tumors grown in 5 mice per condition. (C) Relative expression of macroH2A1 (*H2AFY*) and macroH2A2 (*H2AFY2*) encoding transcripts in resected tumor samples. Rhombi indicate the mean ($n = 10$ tumor grown in 5 mice per condition). A Student's t-test was used to compare the two conditions, * $p < 0.05$. (D) Histological analysis of paraffin-embedded tumor samples by hematoxylin-eosin staining (HE) and immunostaining of Ki67, alpha-fetoprotein (AFP) and β -catenin (β -CAT).

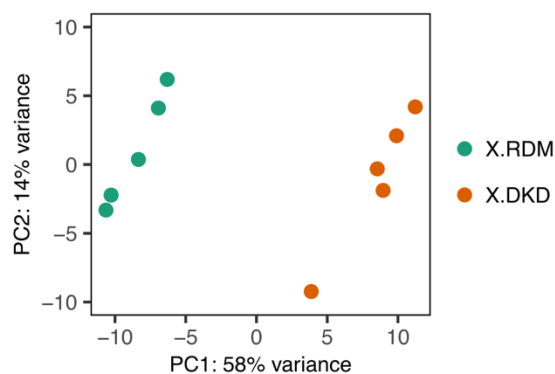
8.3. Transcriptomic profiling of HepG2-derived tumors shows major gene expression changes that depend on macroH2A

We decided to further characterize the xenograft-derived tumors by performing a transcriptomic profiling of gene expression by RNA-Seq. For this we used mRNA extracted from 5 different tumors of each condition. A principal component analysis of the resulting data shows a separate clustering of X.RDM and X.DKD samples along the first principal component (PC1) which represents a high percentage (58%) of the variability in the data (**Figure 39A**). Similarly, hierarchical clustering of the sample-to-sample distances results in two clear clusters that separate X.RDM and X.DKD samples completely (**Figure 39B**).

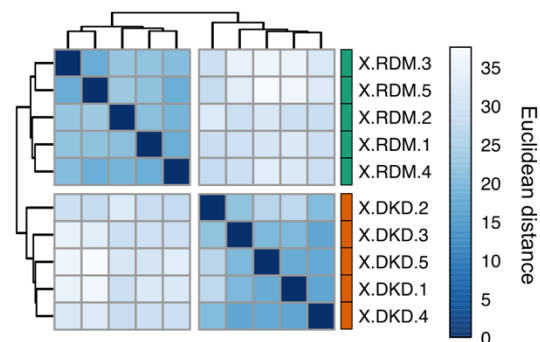
The differential gene expression analysis performed with DESeq2 (Love et al., 2014) reveals over four thousand genes differentially expressed with an adjusted p-value lower than 0.05 when comparing X.DKD versus X.RDM samples with a close to half-and-half proportion of up- and downregulated genes (**Figure 39C**). Of these, 677 genes were differentially expressed with an absolute log₂ fold-change (log₂FC) greater than one (that is, their expression is at least double in the X.DKD group when compared to X.RDM), with a higher proportion of upregulated genes (**Figure 39C**). The subset of genes with an absolute log₂FC ≥ 1 was selected as a relevant dataset for some of the subsequent analysis. The scatterplot of the log₂ mean normalized expression versus the log₂FC (usually referred to as MA plot) does not reveal any systematic bias in the dataset, the differentially expressed genes are well distributed along the normalized expression values and do not come from noisy variation of low expression genes (**Figure 39D**). The scatterplot of log₂FC versus the negative log of the adjusted p-value (usually referred to as Volcano plot) allows the exploration of significantly differentially expressed genes and readily shows the down-regulation of the *H2AFY* and *H2AFY2* genes (**Figure 39E**).

These results shows that we obtained a robust RNA-Seq dataset from tumor samples derived from HepG2 cells which reveals major gene expression differences that depend on macroH2A expression and are not a reflection of growth rate differences.

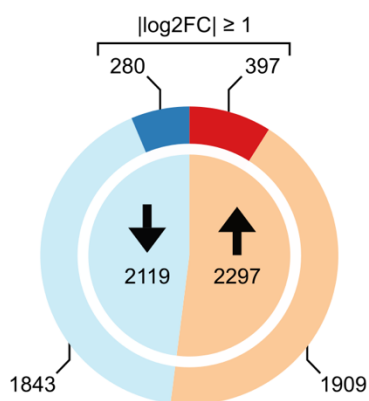
A Principal component analysis



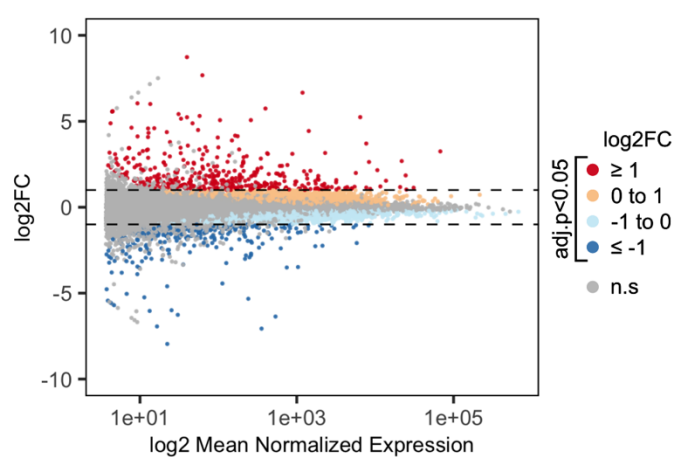
B Sample-to-sample distances



C DE genes adj.p < 0.05 X.DKD vs X.RDM

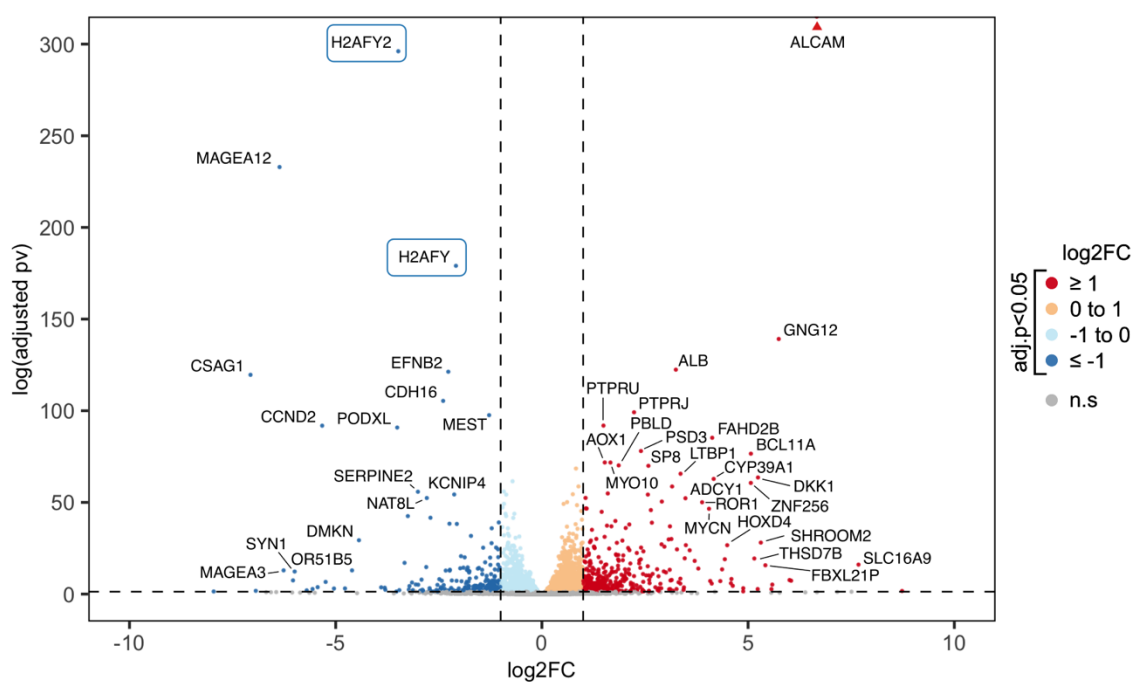


D MA plot X.DKD vs X.RDM



E

Volcano plot X.DKD vs X.RDM



Results

Figure 39 (previous page). Differential gene expression analysis on RNA-Seq data from tumor derived from HepG2 cells. (A) Principal component analysis plot of the 10 sample data included in the study. PC1, principal component 1; PC2, principal component 2. The rlog transformation of the normalized counts was used to calculate the principal components. **(B)** Heatmap of the euclidean distance between samples and hierarchical clustering according to such distances. The white space marks the separation between the two major clusters obtained. The rlog transformation of the normalized counts was used to calculate the euclidean distance matrix between samples. **(C)** Pie chart summarizing the number of differentially expressed genes with an adjusted p-value (adj.p) < 0.05 when comparing X.DKD vs X.RDM. The upwards arrow indicates upregulated genes, while the downwards arrow indicates downregulated genes. The outer ring of the chart further differentiates the subset of genes with an absolute fold change ($|\log_2\text{FC}| \geq 1$). **(D)** MA plot of the \log_2 mean normalized expression versus the \log_2 fold change ($\log_2\text{FC}$) of every gene detected in the data. Dashed horizontal lines mark +1 and -1 $\log_2\text{FC}$. The color code indicates the significance and fold-change range of each gene and follows the same scheme as in (C). **(E)** Volcano plot of the \log_2 fold change ($\log_2\text{FC}$) versus the negative log of the adjusted p-value ($-\log(\text{adj.p})$) of every gene detected in the experiment. The symbols of some low- adj.p and/or high $\log_2\text{FC}$ genes are indicated. The macroH2A1 (H2AFY) and macroH2A2 (H2AFY2) genes are highlighted with a blue frame. The horizontal dashed line marks the significance threshold ($-\log(0.05)$) while the horizontal dashed lines mark +1 and -1 $\log_2\text{FC}$. Color scheme follows the same as in (C) and (D).

8.4. Genes altered in the absence of macroH2A are involved in cellular adhesion, development, hypoxia and inflammation

We next wanted to functionally characterize the genes that showed a differential gene expression dependent on the presence or absence of macroH2A proteins. First, we performed a Gene Ontology (GO) enrichment analysis using only the significant genes with an absolute $\log_2\text{FC}$ equal or greater than 1 (**Figure 39C**) as a way of focusing on the genes that show a bigger effect size in our experiment. The analysis shows various biological process categories significantly enriched in the mentioned gene set, mainly related to cellular adhesion and developmental processes (**Figure 40A**). We grouped these two main GO category groups and extracted the $\log_2\text{FC}$ of the genes included in each category. This shows that there is a mix of up- and downregulated genes in the enriched GO categories with a higher proportion of upregulated genes (**Figure 40B**).

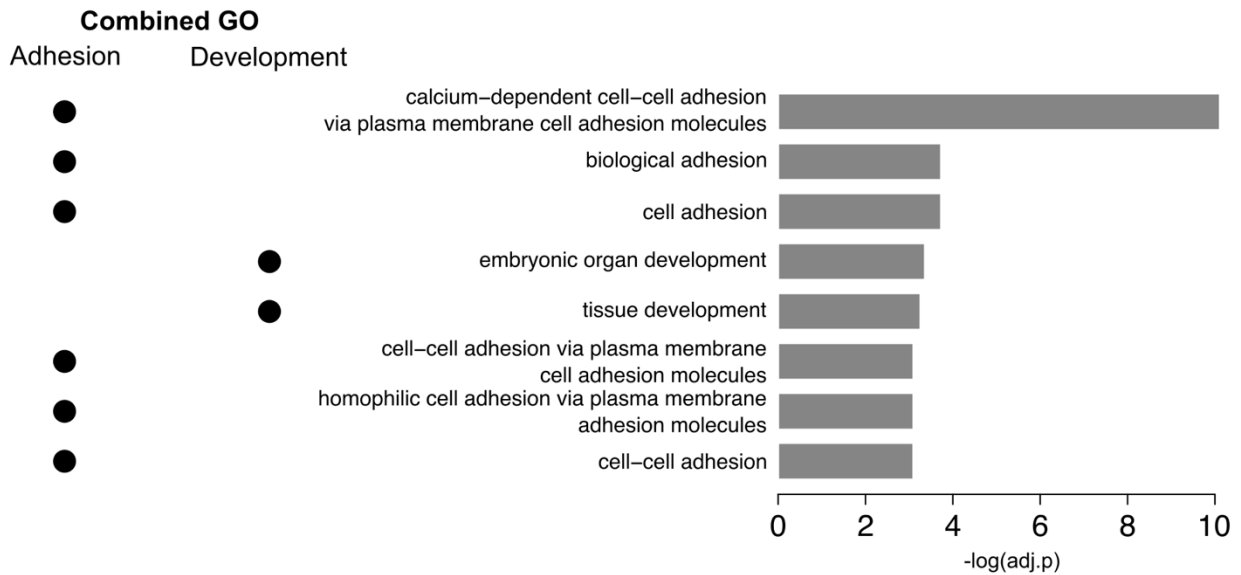
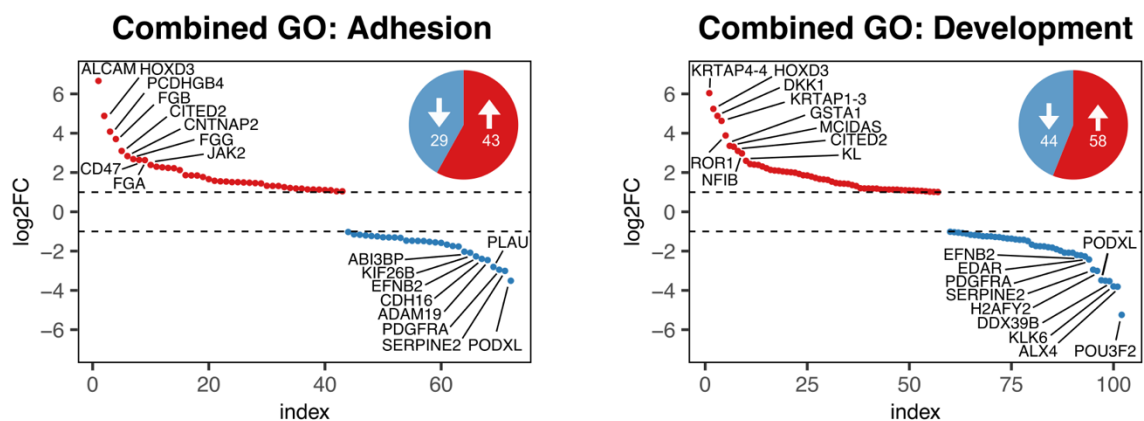
A**Gene Ontology**Input: all DE genes with $|\log_2FC| \geq 1$ **B**

Figure 40. MacroH2A loss deregulates the expression of genes related to adhesion and development. (A) Gene Ontology enrichment analysis using the subset of genes differentially expressed in X.DKD vs X.RDM with an absolute \log_2FC equal or greater than 1. The results are represented as a barplot of the $-\log$ adjusted p-value of each indicated biological process category. On the left side, indication of which related GO categories were combined to simplify the analysis. (B) Plot of the genes pertaining to the “Combined GO” of Adhesion or Development ordered by their \log_2FC . Upregulated genes are colored in red, downregulated genes in blue. The top up- and downregulated genes are annotated. The piechart shows the number of up- and downregulated genes in this category.

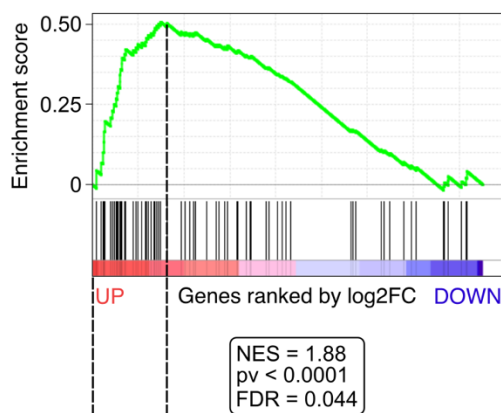
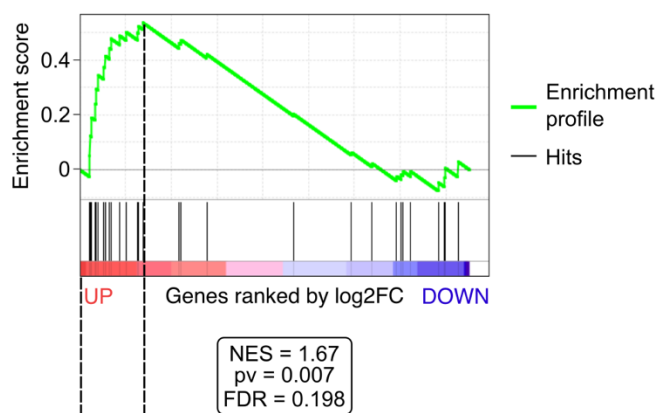
Results

In another approach to identify functional groups of genes deregulated in tumors lacking macroH2A, we used a gene set enrichment analysis using the GSEA software (Mootha et al., 2003; Subramanian et al., 2005). To run the GSEA we used a list of all the genes that were determined as differentially expressed with an adjusted p-value < 0.05 when comparing X.DKD with X.RDM, ranked by their log₂FC value in decreasing order. We did not apply any threshold for log₂FC as we run GSEA with log₂FC as the ranking metric, effectively taking into account the effect size in our experiment. As gene database we selected the hallmark gene sets from the Molecular Signatures Database (MSigDB v6.2 updated July 2018) (Liberzon et al., 2011, 2015) which contains gene expression signatures derived by aggregating many MSigDB gene sets to represent well-defined biological states or processes. The results of this analysis return two gene sets to be significantly enriched and upregulated in our expression dataset. The hypoxia gene set is significantly enriched in genes that become upregulated in tumors lacking macroH2A (**Figure 41A**). Additionally, the inflammatory response gene set is also enriched and upregulated although it has a much higher false discovery rate value (**Figure 41B**). The retrieval of the genes that contribute the most to the enrichment score shows that despite not having filtered the input list for their log₂FC most of them show expression changes with a log₂FC close to 1 (**Figure 41A, B**).

The results of these functional annotation approaches show that the suppression of macroH2A expression in HepG2 cells grown as xenograft tumors in mice provokes the deregulation of genes involved in adhesion and developmental processes, and in particular the upregulation of a set of genes that participate in hypoxia and inflammatory response; all of these biological processes being highly relevant in cancer biology.

A**Gene Set Enrichment Analysis**

Input: all DE genes in X.DKD vs X.RDM with adj.p < 0.05

Hypoxia**Inflammatory response****B**

Gene symbol	log2FC	Running ES
<i>CITED2</i>	3.098	0.0435
<i>NDRG1</i>	2.156	0.0693
<i>LOX</i>	1.991	0.1003
<i>ANGPTL4</i>	1.940	0.1333
<i>PCK1</i>	1.898	0.1660
<i>ERRFI1</i>	1.849	0.1968
<i>F3</i>	1.489	0.2084
<i>PPARGC1A</i>	1.435	0.2313
<i>IGFBP1</i>	1.362	0.2509
<i>DPYSL4</i>	1.265	0.2678
<i>VLDLR</i>	1.219	0.2866
<i>PFKFB3</i>	1.195	0.3066
<i>GPC4</i>	1.138	0.3212
<i>GPC1</i>	1.134	0.3413
<i>SLC2A3</i>	1.131	0.3607
<i>FOS</i>	1.088	0.3759
<i>ADM</i>	1.085	0.3952
<i>PLIN2</i>	1.026	0.4072
<i>NR3C1</i>	1.006	0.4228
<i>PDK1</i>	0.896	0.4192
<i>B4GALNT2</i>	0.863	0.4273
<i>ZFP36</i>	0.840	0.4377
<i>ANKZF1</i>	0.799	0.4416
<i>RBPJ</i>	0.767	0.4458
<i>DTNA</i>	0.761	0.4582
<i>AKAP12</i>	0.743	0.4668
<i>ADORA2B</i>	0.715	0.4716
<i>RORA</i>	0.696	0.4780
<i>PPP1R3C</i>	0.691	0.4878
<i>TGFB1</i>	0.675	0.4950
<i>HDLBP</i>	0.667	0.5021
<i>KLHL24</i>	0.654	0.5079

Gene symbol	log2FC	Running ES
<i>OSMR</i>	2.119	0.0510
<i>IL18R1</i>	2.047	0.1238
<i>LIF</i>	1.929	0.1897
<i>PTGER4</i>	1.661	0.2398
<i>IFITM1</i>	1.581	0.2947
<i>F3</i>	1.489	0.3429
<i>CSF3R</i>	1.257	0.3742
<i>SLC11A2</i>	1.207	0.4148
<i>ADM</i>	1.085	0.4423
<i>SLC1A2</i>	1.052	0.4769
<i>FZD5</i>	0.913	0.4875
<i>ACVR2A</i>	0.833	0.5010
<i>AXL</i>	0.726	0.4985
<i>ADORA2B</i>	0.715	0.5217
<i>IL1R1</i>	0.684	0.5350

Figure 41. Absence of macroH2A leads to the upregulation of genes involved in hypoxia and inflammatory response. (A) Gene set enrichment profiles obtained with the GSEA software using the list of all significantly differentially expressed genes ranked by their log2FC and the hallmark genes sets in the Molecular Signatures Database. NES, normalized enrichment score; FDR, false discover rate; pv, p-value. **(B)** List of core enrichment genes for each of the two gene sets, with their log2FC value. The core enrichment is the subset of genes that contributes most to the enrichment result in the GSEA analysis.

8.5. *DKK1* is upregulated in HepG2-derived tumors lacking macroH2A and hepatoblastoma tumors

To gain further insight into the relevance of the transcriptional changes occurred upon macroH2A downregulation, we crossed the DEG with a $|\log_2FC| \geq 1$ in our experiment with an RNA-Seq expression dataset containing gene expression changes between hepatoblastoma (HB) tumor samples and normal tissue. This dataset was kindly provided by Carolina Armengol (IGTP, Spain). The results show that there is not an overall correlation between the expression changes observed in our experiment with those that happen in hepatoblastoma tumors compared to normal adjacent liver tissue (**Figure 42A**). However, this analysis still allows the identification of a set of genes that change expression in the same sense in the two datasets (**Figure 42A, B**). In particular, there is a strong correlation between the expression change observed in the *DKK1* gene both upon loss of macroH2A and in HB tumors (**Figure 42B**). The *DKK1* gene is, in fact, one of the top up-regulated genes in HB tumors and is the 10th most upregulated gene in our dataset. In addition, *DKK1* mRNA and serum levels are reported to be upregulated in hepatocellular carcinoma and to have value as a prognostic marker (Fouad et al., 2016; Jang et al., 2016; Sakabe et al., 2017). *DKK1* encodes for the Dickkopf-1 (DKK1) protein, a member of the dickkopf family of proteins which are secreted factors that function as a Wnt-signaling inhibitor during vertebrate embryonic development.

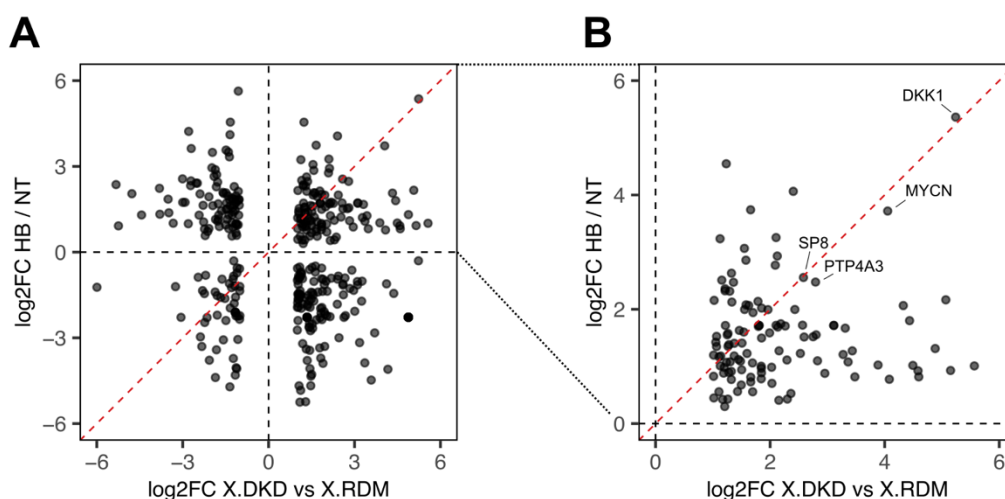


Figure 42. *DKK1* is upregulated in xenografts lacking macroH2A and in hepatoblastoma tumors. (A) Scatterplot of \log_2FC values of the DEG in the contrast X.DKD vs X.RDM that are also significant DEG in an RNA-Seq dataset comparing the expression of hepatoblastoma tumors (HB) and normal liver tissue (NT). Only the genes with a \log_2FC equal or greater than one in the X.DKD vs X.RDM contrast are included in the plot. **(B)** Zoom in of the plot in (A) focused on the genes that are upregulated in both datasets.

Our own ChIP-Seq data in HepG2 cells shows an enrichment of both macroH2A1 and macroH2A2 proximal to the *DKK1* locus (**Figure 43**), which suggests that macroH2A could mediate a direct effect on this gene by altering the chromatin properties in its surroundings. Given the particular relevance of this gene in hepatoblastoma and other cancer types, we chose to use it as a model locus for further study of the role of macroH2A proteins in gene regulation.

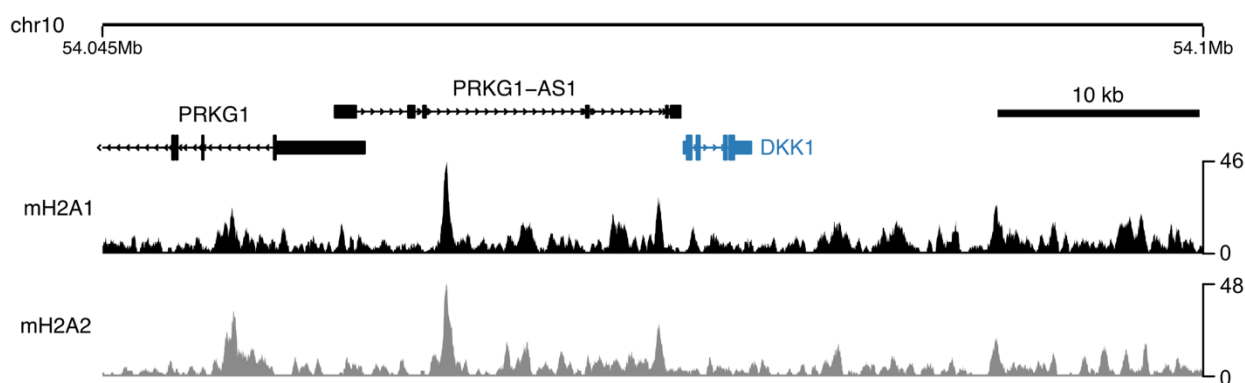


Figure 43. MacroH2A enrichment on the proximity of the *DKK1* locus. ChIP-Seq enrichment profiles for macroH2A1 (black) and macroH2A2 (grey) proximal to the *DKK1* gene locus. On the top, chromosomal coordinates of the area depicted and genes annotated in this region according to the UCSC Genes track.

8.6. The absence of macroH2A results in an upregulation of *DKK1* expression in response to an inflammatory stimulus

We confirmed *DKK1* mRNA levels in the xenograft tumor samples by RT-qPCR, which showed an average of a 20-fold upregulation of this transcript in the X.DKD tumors (**Figure 44A**), quite consistent with the log2FC of 5.2 (36.75 fold-change) detected in the RNA-Seq (**Figure 39E**). Next, we evaluated *DKK1* expression by the same method in the RDM and DKD HepG2 cell lines maintained in cell culture, obtaining a much lower effect close to a 3-fold upregulation (**Figure 44B**). Considering that the environment of the cells when grown in the mice is much more complex than the standard cell culture conditions for HepG2, we thought that some stimulus could be triggering *DKK1* overexpression specifically in the HepG2 cells lacking macroH2A proteins. Interestingly, tumor necrosis factor alpha (TNF α) is reported as a potent inducer of *DKK1* in inflammatory bone disease models (Diarra et al., 2007). This fact, together with the observation that HepG2 xenograft tumors from macroH2A-deficient cells showed alterations in the expression of inflammatory-related genes, led us to hypothesize that an inflammatory stimulus could explain the discrepancy between the expression of *DKK1* in cell culture and xenograft samples.

Results

To test this, we treated HepG2 RDM and DKD cells in cell culture with different concentrations of TNF α and observed a strong upregulation of *DKK1* mRNA abundance specifically in the absence of macroH2A (**Figure 44C**). We could also detect increased levels of secreted DKK1 protein in the culture media of DKD cells compared to RDM cells, with the difference being more prominent upon TNF α treatment in correlation with the RT-qPCR data (**Figure 44D**).

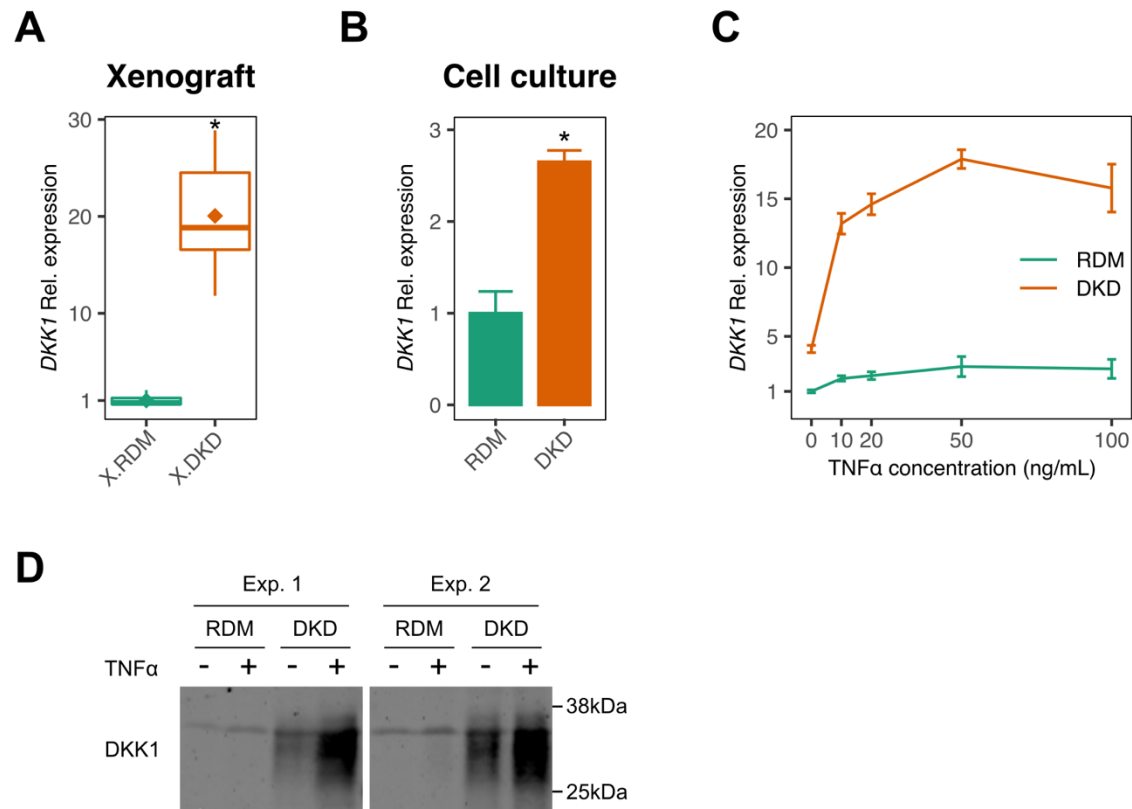


Figure 44. TNF α triggers the overexpression of DKK1 in HepG2 cells lacking macroH2A. (A) Boxplot of the expression of DKK1 in mRNA samples from xenograft tumors derived from HepG2 cell lines. The median is indicated by a horizontal bars and the mean by rhombi. A two-tailed Student's t-test was used to compare the two groups, * p < 0.05 (n = 10). (B) Expression of DKK1 in HepG2 cells in normal cell culture conditions. A two-tailed Student's t-test was used to compare the two groups, * p < 0.05 (n = 3). (C) Dose-response curve of the expression of DKK1 upon treatment HepG2 cells with different concentrations of TNF α for 24h (n = 3). (D) Western blot showing the detection of DKK1 in protein samples precipitated from the culture media of HepG2 cells cultured in the absence (-) or presence (+) of TNF α at 20ng/mL for 24h. Immunoblots of two biological replicates are shown.

8.7. MacroH2A1.1 and macroH2A2 can suppress DKK1 activation in response to TNF α

In order to interrogate if there is any functional difference among the three macroH2A proteins, we generated stable HepG2 cell lines expressing ectopic FLAG-tagged forms of macroH2A1.1, macroH2A1.2 and macroH2A2 in the double macroH2A1 and macroH2A2 DKD background by retroviral transduction. The ectopic constructs were expressed at comparable levels and could recapitulate approximately 50% of the endogenous protein levels in control RDM cells (with the caveat of macroH2A1.1 that is not normally detectable in HepG2) (**Figure 45A**). Immunofluorescence images show that the FLAG-tagged proteins are localized in the nucleus and closely resemble the DAPI DNA staining, indicating that they are incorporated into the chromatin (**Figure 45B**). When treating this panel of cell lines with TNF α , only macroH2A1.1 and macroH2A2 were able to suppress DKK1 activation, while the expression of macroH2A1.2 did not generate a significantly different response in comparison to DKD cells (**Figure 45C**). This result validates the role of macroH2A in repressing *DKK1* expression in HepG2 and highlights the functional specificity of macroH2A1.1 and macroH2A2 in this function.

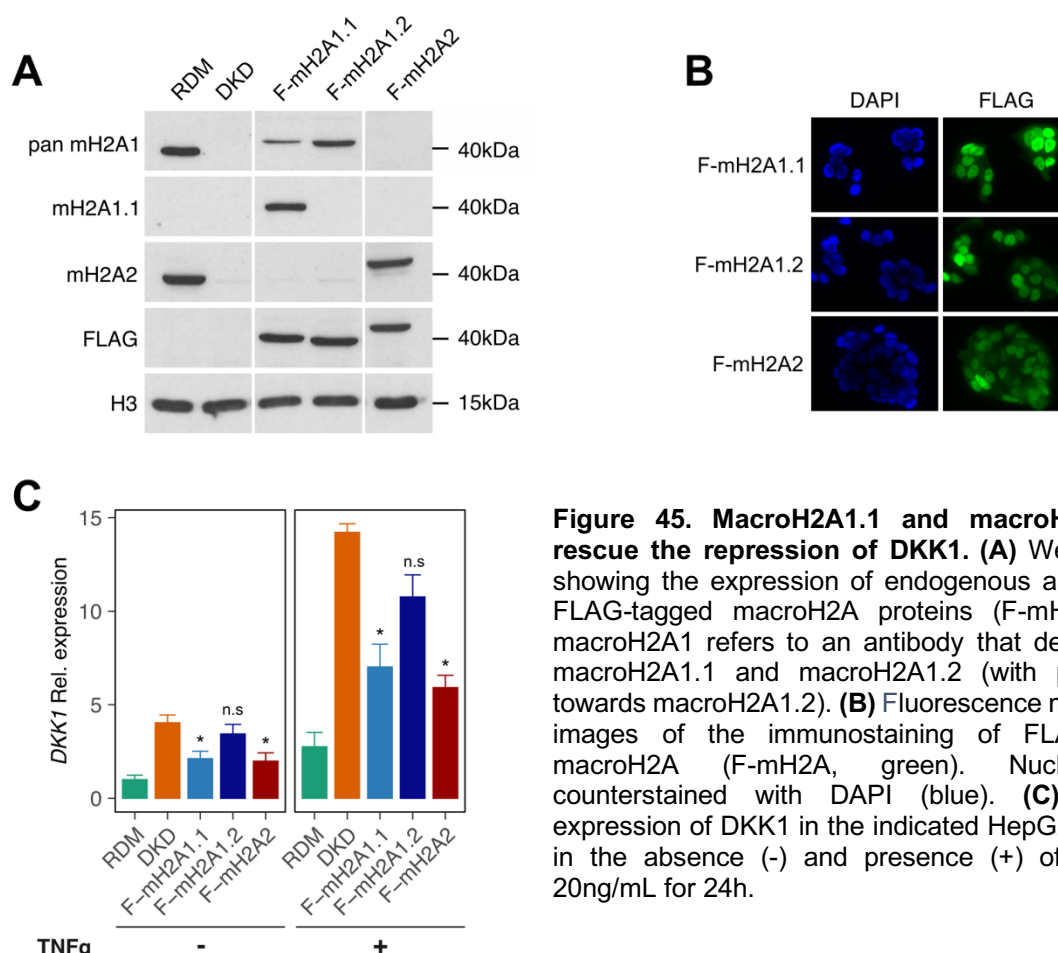


Figure 45. MacroH2A1.1 and macroH2A2 can rescue the repression of DKK1. (A) Western blot showing the expression of endogenous and ectopic FLAG-tagged macroH2A proteins (F-mH2A). Pan macroH2A1 refers to an antibody that detects both macroH2A1.1 and macroH2A1.2 (with preference towards macroH2A1.2). (B) Fluorescence microscopy images of the immunostaining of FLAG-tagged macroH2A (F-mH2A, green). Nuclei were counterstained with DAPI (blue). (C) Relative expression of DKK1 in the indicated HepG2 cell lines in the absence (-) and presence (+) of TNF α at 20ng/mL for 24h.

8.8. Absence of macroH2A alters the 3D structure of the *DKK1* locus and permits a stronger activation of distal regulatory elements upon TNF α treatment

We next sought to study the effect of removing macroH2A in the regulatory landscape of the chromatin at the *DKK1* locus. We used the chromosome conformation capture approach called UMI-4C (Unique Molecular Identifier - 4C, Schwartzman et al., 2016) to quantitatively determine the three-dimensional (3D) contact profile of the *DKK1* promoter in RDM and DKD in standard cell culture conditions (**Figure 46A**). The results show that the *DKK1* promoter has long-range chromatin interactions more prominently towards its 200kb downstream region, which is devoid of genes but contains three elements (E1, E2, E3) annotated as strong enhancers according to their chromatin state (**Figure 46A**). These regions show enrichment in H3 lysine 27 acetylation (H3K27Ac) in ENCODE data and are close to peaks in the UMI-4C signal, which indicates they are contacting the *DKK1* promoter and could act as regulatory elements of this gene (**Figure 46A**). Moreover, the chromatin region spanning E1, E2 and E3 is enriched in macroH2A1 and macroH2A2 ChIP-Seq signal and shows significantly increased contacts with the *DKK1* promoter in DKD cells, indicating that the absence of macroH2A alters the 3D contacts of these enhancers and that it could be due to a direct effect on that particular chromatin region (**Figure 46A**). Additionally, a further downstream region close to a peak of macroH2A shows decreased contacts in DKD cells (**Figure 46A**).

We chose to study if the activity of the E1, E2 and E3 enhancers and two promoter regions close to the *DKK1* transcription start site (P1 and P2, **Figure 46A**) were responsive to the treatment with TNF α and if such response depended on macroH2A. For this, we used chromatin immunoprecipitation to assess the enrichment of these elements in H3 lysine 4 tri-methylation (H3K4me3) and H3 lysine 27 acetylation (H3K27Ac), histone marks mostly associated with active promoters and active enhancers, respectively. The results show that both promoter regions analyzed (P1 and P2) are highly enriched in H3K4me3 while E1, E2 and E3 show much lower levels of this mark (**Figure 46B**). P1 and P2 showed higher levels of H3K4me3 in DKD cells both in the absence and presence of TNF α , indicating that the promoter is more active. Interestingly, the levels of H3K4me3 increased in these two promoter regions when RDM cells were treated with TNF α but in DKD cells the levels of this mark were not changed (**Figure 46B**). Thus, the levels of H3K4me3 in DKD cells did not correlate with the higher *DKK1* mRNA abundance upon TNF α treatment. Similarly, H3K27Ac in the same P1 and P2 regions was more enriched

in DKD cells than in RDM cells (**Figure 46C**). In contrast to H3K4me3, the presence of H3K27Ac at P1 and P2 upon TNF α treatment did not change in RDM cells but slightly increased in DKD cells (**Figure 46C**). We detected H3K27Ac in the three putative enhancers studied, with E3 showing the strongest variation among our experimental conditions. In the absence of TNF α the H3K27Ac signal at E3 was around 2-fold higher in cells lacking macroH2A (**Figure 46C**). The levels of H3K27Ac in this region increased showed a strong activation in DKD cells upon TNF α treatment that brought up the difference between RDM and DKD to almost 4-fold (**Figure 46C**). This E3 element is thus the putative enhancer whose activity (as monitored by H3K27Ac levels) more closely resembles the differential transcriptional response of *DKK1* to TNF α between RDM and DKD cells.

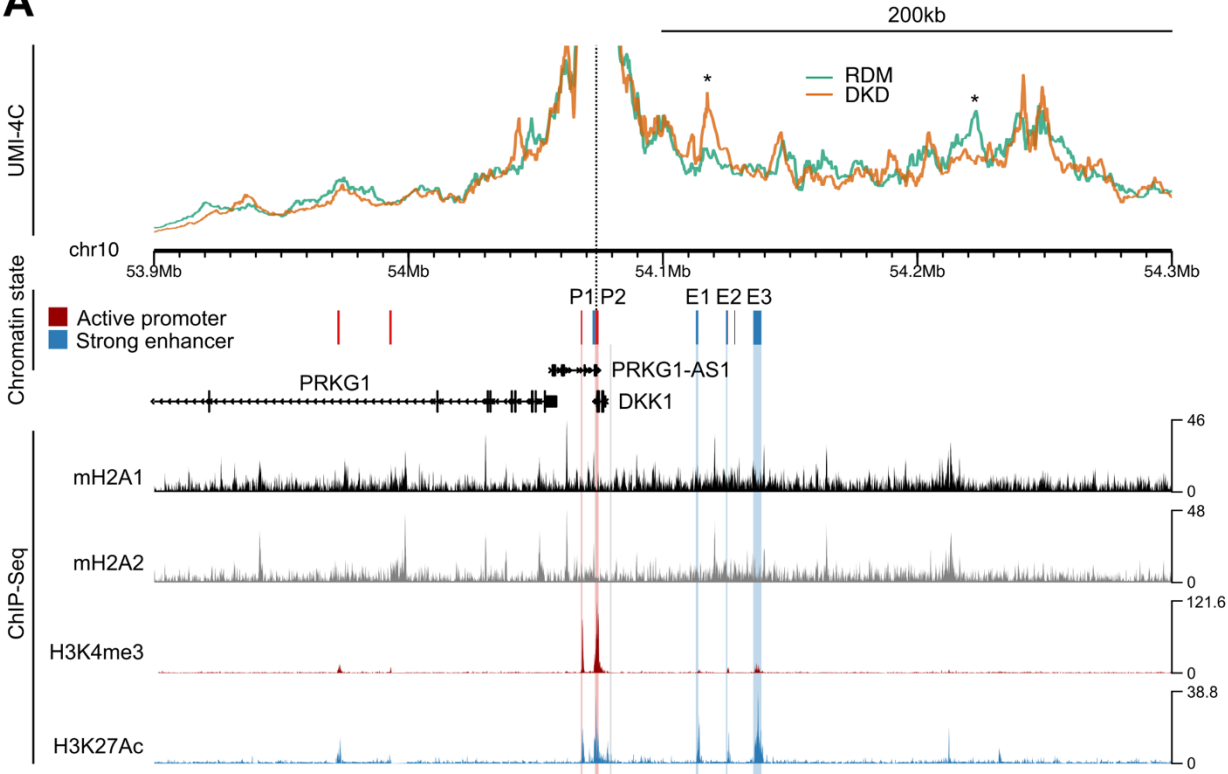
Taken together, these results show that the absence of macroH2A histones alters the 3D contact profile of the *DKK1* promoter with the surrounding regulatory regions as well as the deposition of active histone marks at the *DKK1* promoter and its neighboring enhancers. Additionally, we identified a particular regulatory element annotated as an enhancers that showed a stronger activation in response to TNF α in the absence of macroH2A.

8.9. Contributions

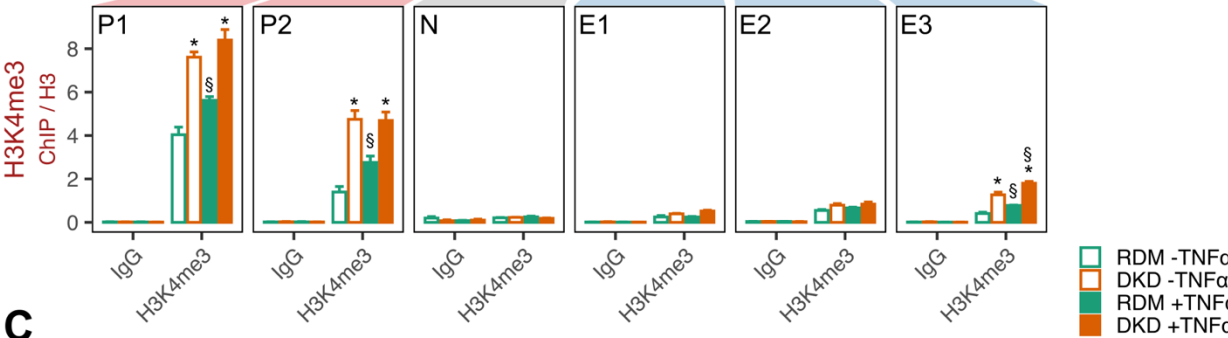
This project was conceived and supervised by Marcus Buschbeck and myself. I generated the stable HepG2 cell lines expressing flag-tagged macroH2A, designed and performed colony formation, wound healing, immunofluorescence, qRT-PCR, western blot and ChIP-qPCR experiments on HepG2 and xenograft samples with great technical support from Vanesa Valero and Ainhoa Pérez. Arce García, Josep Manyé and myself designed and performed the xenograft experiment with technical support from Vanesa Valero and Ainhoa Pérez. I performed the statistical analysis of RNA-Seq data with assistance from Roberto Malinverni. Marguerite-Marie Le Pannérer performed the UMI-4C experiment. Roberto Malinverni pre-processed and mapped RNA-Seq data and analyzed the UMI-4C experiment. Vanesa Valero performed the XTT proliferation assay in HepG2 cells. Laura Royo and Carolina Armengol extracted mRNA from xenograft tumors and performed and analyzed immunohistochemistry of xenograft tumor sections. Carolina Armengol provided the RNA-Seq dataset on hepatoblastoma samples and Juan Carrillo crossed this data with our RNA-Seq on HepG2 xenografts.

Results

A



B



C

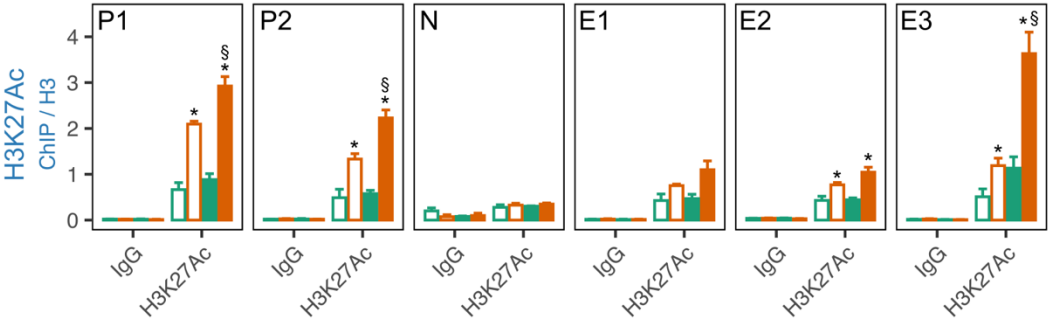


Figure 46. The absence of macroH2A alters the regulatory landscape of DKK1. (A) Chromatin features at the DKK1 locus. From top to bottom: UMI-4C contact profiles of the DKK1 promoter in RDM and DKD cells; chromosome nucleotide coordinates; indicated chromatin states according to the HepG2 Chromatin State Segmentation by HMM from ENCODE and selected chromatin regions as referenced in the main text; our ChIP-Seq data for macroH2A1 and macroH2A2 in HepG2 RDM cells and H3K4me3 and H3K27Ac ChIP-Seq profiles in HepG2 from ENCODE. The dashed horizontal line marks the viewpoint used in the UMI-4C experiment. The UMI-4C profile contains aggregated data from two technical replicates. The * indicate statistically different UMI-4C contact peaks between the two conditions ($p < 0.01$ based on a Chi-Square test as implemented in the Umi4C R package). **(B, C)** ChIP-qPCR data for H3K4me3 and H3K27Ac in the indicated regions shown in (A). The ChIP signal is plotted as the ChIP value for the IP divided by the ChIP of H3, both ChIP values calculated as the percentage of input ($n = 3$). The N region is included as a negative control lowly enriched in both H3K4me3 and H3K27Ac. Cells were treated with TNF α at 20ng/mL for 24h. A two-tailed Student's T-test was used to compare experimental groups, * $p < 0.05$ in the comparison between RDM and DKD in untreated and treated conditions separately, § $p < 0.05$ in the comparison between treated and untreated for each cell line separately.

Discussion

9. Discussion I

9.1. MacroH2A is required for maintaining heterochromatin architecture and attachment to the nuclear lamina

MacroH2A proteins are substantially associated with H3K9me3-marked constitutive heterochromatin at repetitive elements that include pericentromeric satellites and regions of interspersed heterochromatin. This association has largely been underestimated in ChIP-Seq based studies, probably due to the filtering of non-unique sequencing reads in most standard ChIP-Seq analysis pipelines. However, Saksouk *et al.* analyzed by mass-spectrometry the proteins associated with purified pericentromeric major satellites in mouse ESC and found macroH2A1 and macroH2A2 to be enriched at this locus of constitutive heterochromatin (Saksouk *et al.*, 2014). This highlights the limitations of current high-throughput sequencing technologies in dealing with repetitive regions and the value of alternative locus-purification methods.

The loss of macroH2A proteins in HepG2 cells resulted in major changes in the organization of constitutive heterochromatin. Repetitive sequences were disorganized, partially decondensed and fragmented. This effect correlated with a loss of binding to lamin B1. Wen *et al.* have reported an enrichment of macroH2A in biochemical purifications of lamin B1 (Fu *et al.*, 2015). We could extend this finding by showing that both macroH2A1 and macroH2A2 interact with lamin B1 at the nuclear periphery in human HepG2 cells. Moreover, we show that macroH2A is required for the interaction of specific repetitive elements with lamin B1. Wen *et al.* also observed a loss of the bright dense signals of H3K9me3 foci in mouse cells upon knock-down of macroH2A1. We found a similar phenotype of dispersed H3K9me3 staining in HepG2 and could additionally demonstrate that removing macroH2A results in a global reduction in the interactions detected between H3K9me3-marked chromatin and lamin B1.

9.2. Possible mechanisms of macroH2A-lamin B1 interaction

The presence of H3K9me3 is required for the tethering of constitutive heterochromatin to the nuclear lamina in human and *C. elegans* (Bian *et al.*, 2013; Gonzalez-Sandoval *et al.*, 2015; Harr *et al.*, 2015; Towbin *et al.*, 2012). Our results add macroH2A as a novel factor contributing to this interaction without being required for H3K9me3, as removal of macroH2A proteins did not affect the levels of H3K9me3 at

Discussion

repetitive elements. It is currently not clear whether the interaction between macroH2A and lamin B1 is direct or indirect. Both our results and Wen *et al.* study used proximity-based methods that can detect interactions happening at 20-30nm, which would be compatible with an indirect interaction mediated by a third factor. Many inner nuclear membrane proteins have been implicated in the attachment of chromatin to the nuclear lamina, and examples of direct and indirect interactions have been described (**Figure 14**). For instance, the lamin B receptor (LBR) is an inner nuclear membrane protein that attaches chromatin to the nuclear lamina through binding of H4K20me3 or via an interaction with HP1 (Hirano et al., 2012; Polioudaki et al., 2001; Ye et al., 1997). The presence of macroH2A could promote the function of other factors or provide an additional interaction surface through its macrodomain.

Single-cell mapping of LADs using DamID suggests a model of multivalent interactions of long stretches of the genome in continuous contact with the nuclear lamina rather than the existence of focal interaction points (Kind et al., 2015). Our results show that macroH2A does not affect the compaction of the chromatin fiber in a one-dimensional manner, which suggests that the alterations in heterochromatin architecture that we observed are due to loss of higher-order structures. The maintenance of such higher-order structures could be an important factor for keeping the local concentration of multivalent interactions sites sufficiently high for a stable interaction of long chromatin regions with the nuclear lamina (**Figure 47**). This mechanism could explain the reorganization of heterochromatin observed during senescence. Cellular senescence is characterized by a progressive loss of lamin B1 proteins (Freund et al., 2012; Shah et al., 2013) and the accumulation of macroH2A-enriched dense heterochromatin foci in the nucleosol (Zhang et al., 2005b). This could reflect that upon loss of lamin B1-mediated interactions at the nuclear periphery, macroH2A is still involved in maintaining dense heterochromatic structures in the nuclear interior.

9.3. MacroH2A is not essential for the silencing of repetitive elements

Removal of macroH2A provoked architectural changes and loss of lamina interactions of several repetitive elements, but this only resulted in a two-fold expression change for the heterochromatic repeats *SAT2* and *DXZ4* which have a very low endogenous expression level. Chromatin accessibility, nucleosome density and H3K9me3 levels at these loci did not change, suggesting that the main features of transcriptional repression were not affected. This indicates that macroH2A is involved in

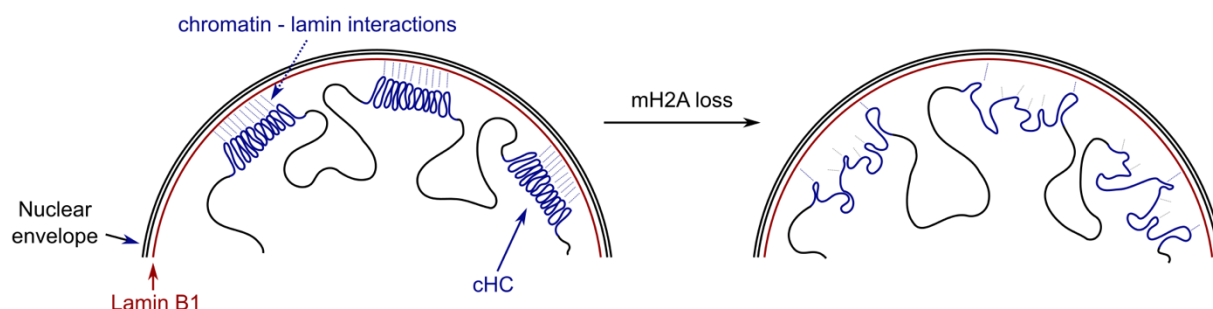


Figure 47. Model for macroH2A-promoted lamina attachment. The maintenance of higher-order heterochromatic structures could be an important factor for keeping a high local concentration of multivalent interactions sites necessary for a stable interaction of long chromatin regions with the nuclear lamina.

the higher-order organization of H3K9me₃-marked heterochromatin but is not an essential factor for the transcriptional repression of these regions. In the course of our immunofluorescence experiments, we did not observe evidences of chromosomal instability or mitotic defects indicative of genomic instability, which is consistent with the silent state of repetitive elements. However, we did not conduct any specific experiment or analysis to evaluate genomic instability which would be needed to make a statement in this regard. The constitutive heterochromatin compartment has highly redundant silencing mechanisms and in most cases the removal of a single factor does not completely disrupt repression. For instance, in the inactive X chromosome (Xi), macroH2A is incorporated late in the inactivation process but is not required for X inactivation (Mermoud et al., 1999; Tanasijevic and Rasmussen, 2011). However, the presence of macroH2A1 prevents Xi reactivation in the presence of DNA methylation and histone deacetylation inhibitors (Hernández-Muñoz et al., 2005). The role of macroH2A in the silencing of repetitive elements could be similar, where it is not necessary for repression in steady-state conditions but provides an additional layer of stability relevant during challenges.

9.4. MacroH2A maintains nuclear and nucleolar integrity

Apart from the reorganization of heterochromatic repeats, we observed a higher proportion of nuclei with irregular shape and larger volume upon removing macroH2A. The nuclear lamina and chromatin are major mechanical components in determining the shape of the nucleus (Stephens et al., 2019). Reduced chromatin compaction forced by treatment with HDAC inhibitors and histone methyltransferase inhibitors induces nuclear defects in the form of blebs without perturbing the nuclear lamina filaments (Stephens et al., 2018). Similarly, it is conceivable that macroH2A-dependent changes in

Discussion

heterochromatin organization, in particular in the nuclear periphery, affect the shape and size of the nucleus

In addition, in macroH2A deficient cells the nucleolus was expanded and had lost the heterochromatin shell at its surface according to electron microscopy images. The nuclear lamina and the nucleolar surface both accumulate constitutive heterochromatin and contain similar repetitive sequences (Koningsbruggen et al., 2010; Németh et al., 2010b). After cell division, loci associated to the nuclear lamina can shuttle to the nucleolar surface in the daughter cell (Kind et al., 2013a). Interestingly, lamin B1 is necessary for maintaining the structural plasticity of nucleoli (Martin et al., 2009) and lamin B1 loss during senescence correlated with impaired interactions between the perinucleolar space and satellite repeats (Dillinger et al., 2017). Thus, macroH2A is a factor that maintains the organization of constitutive heterochromatin in the nuclear and nucleolar periphery, which are spatially different but similar repressive compartments.

Loss of macroH2A resulted in a significant increase in the expression of rDNA, which was also observed in a previous study when removing only macroH2A1 (Cong et al., 2014). Considering that ribosomal RNAs are the most abundant RNA molecules in the cell, it can be considered the most drastic effect of macroH2A on the expression of repeats. rDNA is organized in clusters where a fraction of repeat units are kept silent in the form of H3K9me2/3-marked heterochromatin (Zillner et al., 2015). The three-dimensional organization of the clusters in the nucleolar environment is believed to be important for their transcriptional activity. Accessibility assays in a previous study suggested that the number of active rDNA units is not affected by the absence of macroH2A1 (Cong et al., 2014), similar to our observations when removing both macroH2A1 and macroH2A2. This further supports a role for macroH2A in higher-order organization of repeats rather than a direct transcriptional control.

9.5. The function of macroH2A in higher-order structures may not be limited to constitutive heterochromatin

MacroH2A proteins are found in other chromatin environments that are not constitutive heterochromatin. Genomic profiling studies, including our own data, have identified a prominent association of macroH2A with H3K27me3-marked polycomb domains. In this context macroH2A is found covering the promoter and gene body of silenced genes, following a very similar distribution to H3K27me3 (Buschbeck et al., 2009; Gamble et al., 2010; Gaspar-Maia et al., 2013). Polycomb domains share some

characteristics with constitutive heterochromatin, mainly a transcriptional repression that is thought to be mainly mediated by chromatin compaction. In addition, microscopy and chromosome conformation capture studies have revealed the existence of long-range interactions between polycomb-repressed chromatin. It is tempting to speculate that macroH2A could participate in maintaining higher-order structures in the three-dimensional organization of polycomb domains (Entrevaan et al., 2016; Ogiyama et al., 2018).

10. Discussion II

10.1. MacroH2A histone variants show diversity in structure, expression and function

The three proteins that form the macroH2A family of histone variants in vertebrates share a conserved domain structure. However, their sequence differences impact the structure of the macrodomain and in particular the properties of its pocket (Posavec et al., 2013). MacroH2A1.1 is the only isoform so far for which a ligand to the macrodomain has been identified. MacroH2A1.1 can bind ADP-ribose while macroH2A1.2 and macroH2A2 are unable to do so. Interestingly, the macrodomain sequence of macroH2A2 is also highly conserved. A potential binding function of macroH2A2 could explain this conservation, raising the possibility that the ligand for this macrodomain remains yet to be described.

Our results and others reveal a differential tissue expression of macroH2A proteins (Pehrson et al., 1997), and many examples of functional differences among macroH2A proteins have been reported. Most notably, macroH2A1.1 expression generally correlates with differentiation while macroH2A1.2 is the predominant form in undifferentiated and cancer cells (Posavec-Marjanović et al., 2017; Sporn and Jung, 2012). High macroH2A1.2 to macroH2A1.1 ratio increases migration, invasion and growth of breast cancer cell lines (Dardenne et al., 2012; Li et al., 2012). The process of somatic cell reprogramming is another example of functional diversity, where macroH2A2 is the main form acting as a barrier to induced pluripotency (Gaspar-Maia et al., 2013). In contrast, the genomic enrichment of different macroH2A proteins in bulk cell samples shows a high overlap (Gaspar-Maia et al., 2013; Pehrson et al., 2014). This raises the question of whether relative abundance is the sole determinant of the proportion of different macroH2As found at enriched genomic sites.

10.2. MacroH2A1.1 can regulate PARP1-dependent processes

Our study shows that macroH2A1.1 can inhibit chromatin rearrangements by inhibiting PARP1. MacroH2A1.1 binds and inhibits PARP1 in its ADP-ribosylated form in a manner that is fully dependent on the structure of its macrodomain binding pocket (**Figure 48A**) (Posavec-Marjanović et al., 2017; Timinszky et al., 2009). A previous study reported the capacity of all macroH2A proteins to inhibit PARP1 (Nusinow et al., 2007b). We demonstrate that PARP1 inhibition is specific to macroH2A1.1 and is not found in macroH2A1.2 or macroH2A2.

Discussion

This specific function of macroH2A1.1 puts functional weight on the splicing event of the macroH2A1 transcript. During myogenic differentiation, this splicing event switches towards increased macroH2A1.1 expression (Posavec-Marjanović et al., 2017). This leads to a reduced consumption of NAD⁺ by PARP1 in the nucleus, which favors mitochondrial NAD⁺ salvage and respiration. MacroH2A1.1 binding to PARP1 is also implicated in gene regulation. In HeLa cells subject to heat-shock, macroH2A1.1-mediated PARP1 recruitment is required for the expression of stress-response genes (Ouararhni et al., 2006). In fibroblasts, PARP1 binds and cooperates with macroH2A1.1 to both positively and negatively regulate gene expression (Chen et al., 2014). Our results show that the inhibitory effect of macroH2A1.1 in PARP1-dependent chromatin relaxation is partial when compared to that of a potent small-molecule inhibitor. Taken together, this observations suggest that the presence of macroH2A1.1 may set a threshold for PARP1 activation. In this way, the intensity of stress signals could determine if macroH2A1.1 is able to globally inhibit PARP1 activation or rather acts as a recruiting factor for PARP1 to regulated genes.

10.3. Possible mechanisms of stabilization of chromatin structures by the macroH2A linker

Our study has demonstrated that the linker region of all macroH2A proteins is able to limit chromatin expansion in response to acute DNA-damage. Remarkably, the addition of the macroH2A2 linker to the canonical histone H2A was sufficient to confer this function. In addition, the linker of macroH2A2 is sufficient to rescue the defective organization of H3K9me3-marked heterochromatic structures observed upon loss of macroH2A. However, the full-length protein was slightly more effective than the histone-fold and linker alone, which suggests that the macrodomain also contributes to this function, possibly by providing an additional surface that could mediate protein-protein interactions.

The linker domain of macroH2A consists of about 40 amino acids that connect the histone fold with the macrodomain. This linker is highly flexible and is excluded from all available crystallographic structures of macroH2A (Chakravarthy et al., 2005b; Kustatscher et al., 2005). It is considered to be an intrinsically disordered region and contains very few order-inducing amino acids with a high percentage of basic, positively charged residues (Muthurajan et al., 2011). Although the linker shows a very low conservation in the alignment with zebrafish and medaka proteins, the disorder-prone nature and high percentage of lysines and arginines is a conserved characteristic of this

stretch of the proteins (**Figure 48B**). This features resemble the C-terminal tail of the linker histone H1 (Caterino et al., 2011). Interestingly, H1 and macroH2A have a mutually exclusive relationship in chicken chromatin (Abbott et al., 2005) and a cross-reactivity of anti-macroH2A antibodies with H1 has been reported in two instances (Pehrson et al., 1997; Rivera-Casas et al., 2016). In the case of H1, the intrinsic disorder of the C-terminal region and its charge-biased amino acid composition are essential for its function in compacting chromatin, rather than the precise sequence (Lu et al., 2009). Interestingly, the PRC1 subunit Cbx2 also contains a region of high positive charge that is necessary for compaction of nucleosomal arrays *in vitro* (Grau et al., 2011). It would be informative to study if this loose sequence requirements are similar for the function of the macroH2A linker by testing the effect of shuffling the position of residues without affecting the overall amino acid composition. Additionally, keeping the flexibility of the linker but neutralizing the positive charges with point mutations or switching them to acidic residues should result in a loss of function if the proposed mechanism is correct.

Pioneering *in vitro* experiments showed that the linker of macroH2A1 possesses properties similar to those of H1. The linker stabilizes DNA at the entry/exit site of the nucleosome, reducing its accessibility to DNA exonucleases (Chakravarthy et al., 2012). In addition, the macroH2A1 linker enhances the condensation of nucleosomal arrays and promotes fiber-fiber interactions *in vitro*, however in the absence of the macrodomain exclusively (Muthurajan et al., 2011). We hypothesize that the capacity of the linker to generally compact chromatin fibers *in vitro* could explain its ability to reduce chromatin relaxation upon DNA-damage without interfering with PARP1 and promote the organization of constitutive heterochromatin. The linker properties might also explain how macroH2A1.2 contributes to the compaction of chromatin at a later time point of the double-strand repair process (Khurana et al., 2014). The mediation of fiber-fiber interactions by the macroH2A linker could explain the architectural changes observed in repetitive elements without changing the nucleosomal density and the one-dimensional chromatin compaction.

In support of a role of the macroH2A linker in chromatin compaction, *Arabidopsis thaliana* lacks macroH2A proteins but contains H2AW histone variants with structural and functional similarities with macroH2A. H2AW contains an extended C-terminus that resembles part of the linker region of macroH2A and the C-terminus of H1. This region is required for heterochromatin condensation *in vivo* and promotes the self-association of nucleosomal arrays *in vitro* (Yelagandula et al., 2014).

Discussion

In conclusion, the biophysical properties of the linker of macroH2A are similar to those of the C-terminal tail of H1 and are compatible with the mediation of fiber-fiber interactions through electrostatic interactions that is supported by *in vitro* experiments. This provides an hypothetical mechanism that can explain the function of macroH2A in heterochromatin organization, promotion of lamina attachment and suppression of chromatin rearrangements in response to DNA damage.

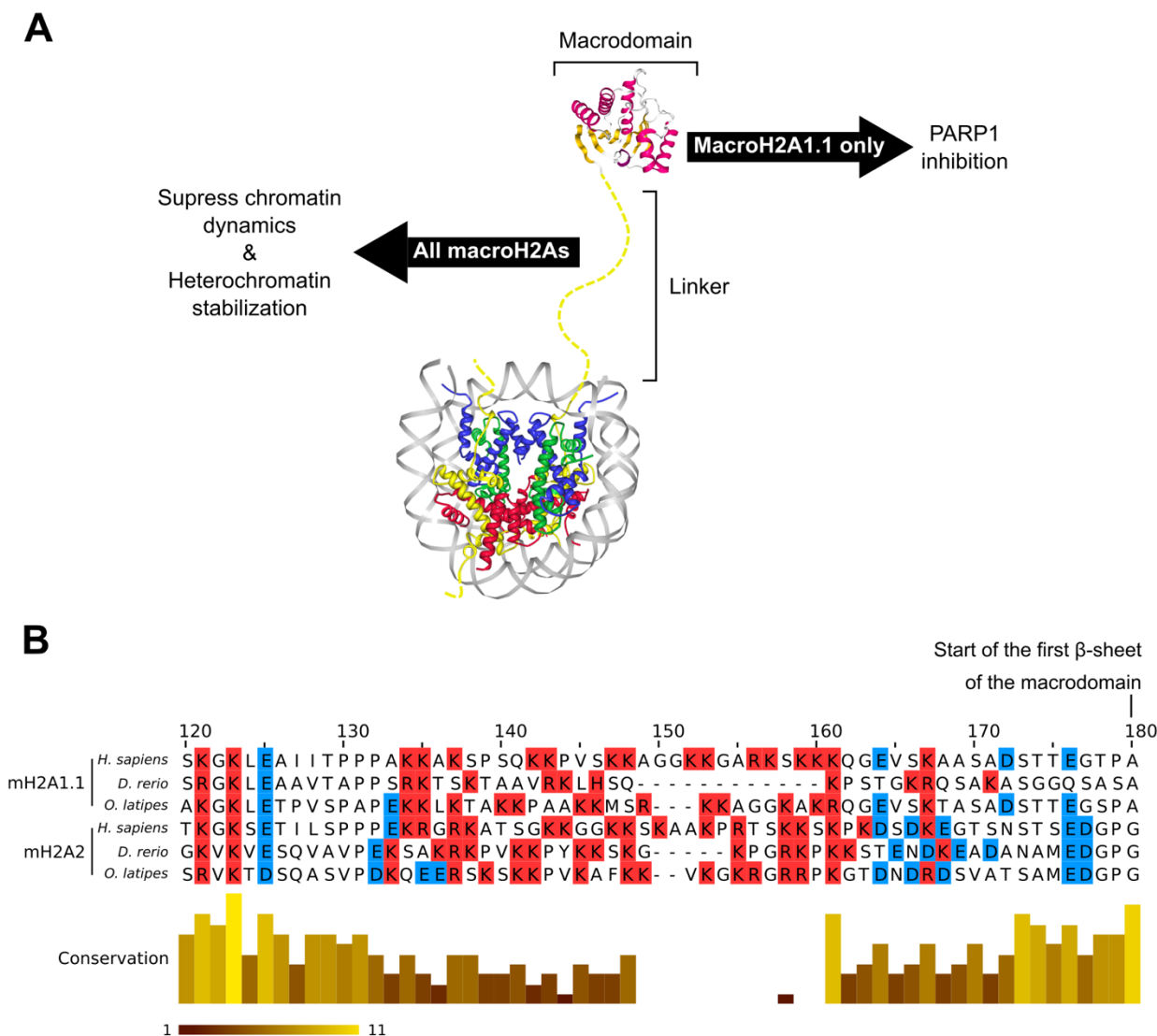


Figure 48. MacroH2A histones can affect chromatin organization through two distinct mechanisms. (A) Schematic summary of the proposed model. The macrodomain of macroH2A1.1 can bind ADP-ribose and inhibit PARP1, suppressing PARP1-mediated processes. At the same time, the linker region of all macroH2A can generally suppress chromatin dynamics and participates in heterochromatin stabilization. (B) Multiple sequence alignment of human (*H. sapiens*), zebrafish (*D. rerio*) and medaka (*O. latipes*), only the sequence around the linker region is shown. Basic, positively charged amino acids are highlighted in red, while acidic, negatively charged amino acids are highlighted in blue. Most studies consider the linker to extend from residues 123 up to 161, although the first element of the macrodomain with a defined secondary structure starts at residue 180. Amino acid coordinates refer to the sequence of human macroH2A1.1 (and coincide with human macroH2A2), the alignment is the same as the one shown in Figure 28.

10.4. Potential regulation of the macroH2A linker by PTMs

The linker shared by both macroH2A1 splice isoforms has been reported to be phosphorylated on T128 and S137 (Bernstein et al., 2008; Nusinow et al., 2007b). T128 is replaced by a serine in macroH2A2 but its potential phosphorylation has not been analyzed so far. On the contrary, S137 is replaced in macroH2A2 by a lysine which can not be phosphorylated. Knowledge on the function of these modifications remains scarce, but they pose the question of whether the function of the linker could be regulated by the presence of phosphorylated residues. S137-phosphorylated macroH2A1 is excluded from the inactive X chromosome and, while only a small fraction of macroH2A1 is phosphorylated, its abundance greatly increases during mitosis (Bernstein et al., 2008). Both Haspin kinases as well as cell cycle regulators Cdk1/cyclinB and Cdk2/cyclinE have been suggested to catalyze the S137 phosphorylation (Bernstein et al., 2008; Maiolica et al., 2014). We hypothesize that the addition of negative charges to the macroH2A linker would reduce its ability to maintain chromatin interactions by neutralizing its positively charged nature. The timing of the phosphorylation of macroH2A at S137 resembles that of H1, which also becomes hyperphosphorylated during mitosis by Cdk2. Phosphorylation of the C-terminal domain of H1 is thought to destabilize H1-chromatin interactions and promote chromatin relaxation *in vitro* and *in vivo* (Contreras et al., 2003; Lopez et al., 2015).

ADP-ribosylation is a post translational modification that has a great capacity to confer multiple negative charges. Considering the tight relation and colocalization of macroH2A1.1 and PARP1 in some chromatin contexts, it is possible that when the inhibition of PARP1 is overcome by stress, macroH2A itself is ADP-ribosylated which could impair its interactions with chromatin.

The fact that the phosphorylation of the macroH2A linker happens in a specific time frame highlights the dynamic nature of PTM-mediated regulation. It is possible that mass-spectrometry experiments have underestimated or missed the presence of macroH2A PTMs that happen only in a subpopulation of cells which are undergoing a certain cell cycle stage or are under a particular stimulus, for instance. Histone H1 phosphorylation by Cdk2, for example, is associated with transcriptional activation in response to hormones (Bhattacharjee et al., 2001; Vicent et al., 2011). It would be interesting to analyze the possible PTMs of macroH2A proteins in synchronized or stimulus-challenged cell cultures to try to reveal if other time-restricted or signal-induced PTMs exist.

11. Discussion III

11.1. MacroH2As are context- and isoform- dependent tumor suppressors that impact on transcriptional regulation

MacroH2A histones are primarily considered to act as tumor suppressors, although context-dependent and isoform-specific scenarios are common (2.2.5). Our results show that reducing the expression of all macroH2A proteins in HepG2 cells does not affect their growth rate in culture but increases their migratory capacity. A recent study reports that macroH2A1.1 overexpression, but not macroH2A1.2, can block epithelial-to-mesenchymal transition (EMT) in mammary epithelial human cells (Hodge et al., 2018). HepG2 cells express very low levels of macroH2A1.1, which makes this protein highly unlikely to be responsible for this phenotype in HepG2. On the other hand, macroH2A1.2 inhibits the converse process of mesenchymal-to-epithelial transition in mesoderm-derived human fibroblasts subject to reprogramming (Pliatska et al., 2018). Testing the expression of EMT markers could reveal if the increased migratory phenotype is related with an EMT process. In addition, the use of unique macroH2A knock-downs and rescue cell lines could reveal which is the role of each macroH2A protein in the migratory phenotype observed in HepG2.

Suppression of macroH2A expression in HepG2 did not alter their growth as subcutaneous xenografts in immunosuppressed mice nor did it provoke major histological differences in the tumor tissue. A genome-wide transcriptional analysis revealed significant differences in gene expression with more than 600 genes de-regulated with an absolute fold-change greater than 2. This gene set contained a comparable number of up- and downregulated genes. Together with many other observations, this challenges the role of macroH2A histones as sole transcriptional repressors (Barrero et al., 2013a; Chen et al., 2014; Creppe et al., 2012; Gamble et al., 2010; Ouararhni et al., 2006). Functional analysis of the deregulated genes revealed an enrichment in genes related to cell adhesion, developmental processes, hypoxia and inflammation, all of which are processes relevant for cancer progression (Hanahan and Weinberg, 2011). All these subsets of genes contained both up- and downregulated genes, and a detailed pathway-based analysis will be needed to unravel how the transcriptional alterations relate to cellular phenotypes. The observed changes in gene expression are probably a combination of events that depend more directly on macroH2A and others which are

Discussion

secondary to altered transcriptional regulation pathways. A possible approach to unravel this would imply integrating the ChIP-Seq enrichment of macroH2A with the position of the DEGs identified, for example selecting the DEGs that are within a certain distance of macroH2A peaks and enriched domains. Importantly, the transcriptional deregulation observed is not derived from different cellular proliferation rates, which is a major confounding factor when analyzing DEGs in response to treatments.

We focused our attention on *DKK1*, a gene that was upregulated in xenografts lacking macroH2A and in hepatoblastoma tumor samples compared to normal tissue. Our ChIP-Seq data showed proximal enrichment of macroH2A to the *DKK1* locus, which suggested that it could be a gene directly regulated by the presence of macroH2A. The repression of *DKK1* could be rescued by macroH2A1.1 and macroH2A2, but not macroH2A1.2 despite being the most abundant macroH2A protein in HepG2 cells. MacroH2A1.1 is the macroH2A histone most clearly correlated with differentiated states and tumor suppression, while macroH2A1.2 is highly expressed in undifferentiated cancer cell lines (Sporn and Jung, 2012). However macroH2A1.1 is lowly expressed in HepG2, which indicates that endogenous macroH2A2 is likely to be responsible for the repression of *DKK1*. This observation provides yet another example of the functional diversity among the macroH2A family of histone variants.

There are examples where macroH2A2 was shown to be the most relevant macroH2A functionally despite being usually expressed at lower levels. For instance, macroH2A2 is the most effective in blocking reprogramming induced by Yamanaka factors in MEF (Gaspar-Maia et al., 2013) while knocking down mH2A2 in melanoma increased its metastatic potential much more than knocking down mH2A1 (Kapoor et al., 2010). ChIP-Seq data, however, reveals an almost identical distribution of macroH2A1.2 and macroH2A2, so their unique functions are likely to rely on the difference in structure and potential interactors rather than a specific genomic localization.

11.2. MacroH2A represses *DKK1*, a gene upregulated in hepatoblastoma and liver cancer

Hepatoblastoma (HB) is a pediatric tumor of hepatocellular origin that is usually diagnosed in children before 5 years of age (Spector and Birch, 2012). It represents around 1% of all pediatric malignancies and is one of the most common malignant liver pediatric neoplasms (Zimmermann, 2005). It is, however, a very rare disease with an incidence of 1.5 cases/million children under 15 years in western countries, with a slightly

higher incidence in males (Schnater et al., 2003; Spector and Birch, 2012). The HepG2 cell line was established in 1979 from a liver tumor of a 15-year old caucasian male and, although initially it was mistakenly reported as a hepatocellular carcinoma cell line, it is currently validated as an hepatoblastoma-derived cell line and used to model hepatoblastoma and liver functions in several studies (López-Terrada et al., 2009a).

HB is thought to have an embryonal origin and arise from an hepatocyte precursor (Zimmermann, 2005). The histopathology of HB is heterogeneous and usually reproduces different stages of liver development with a mix of histological profiles and subtypes that include epithelial, mesenchymal and undifferentiated characteristics (Cairo et al., 2008; Tanaka et al., 2013; Zimmermann, 2005). The histological analysis of the tumors derived from HepG2 cells in our xenograft experiment showed a prevalence of undifferentiated tissue and a pleomorphic pattern with high mitotic activity. The pleomorphic epithelial pattern is uncommon in HB and contains cells with fetal/embryonal appearance (Tanaka et al., 2013). No alterations on the histological properties or the expression of pathological markers was observed upon removal of macroH2A.

The transcriptional changes dependent on macroH2A observed in HepG2 did not show an overall correlation with the differences between HB tumors and normal tissue. In fact, both macroH2A1 and macroH2A2 are actually slightly upregulated in HB (log2FC 0.83 and 1.38, respectively), which indicates that the disease scenario does not mimic our approach of suppressing macroH2A expression in HepG2. *DKK1* encodes for Dickkopf-1 (DKK1), a secreted protein member of the Dickkopf family that is mainly characterized as an inhibitor of β -catenin-dependent Wnt signaling (Bafico et al., 2001; Mao et al., 2001). *DKK1* stood out as being highly upregulated (almost 6-fold) in both datasets. Deletion mutations stabilizing β -catenin are prevalent in hepatoblastoma and hepatocellular carcinoma and have also been characterized in HepG2 cells (Fatima, 2011; López-Terrada et al., 2009b). Wnt signaling in HB tumors promotes the expression of markers of hepatic progenitors (Cairo et al., 2008). The overexpression of DKK1 in HB is thought to be an attempt to activate a negative feedback to balance Wnt signaling (Cairo et al., 2008). However, DKK1 has also been implicated in the activation of β -catenin-independent Wnt signaling (Caneparo et al., 2007; Endo et al., 2008; Pandur et al., 2002). The HepG2 xenografts we generated were positive for β -catenin staining and we did not observe changes on β -catenin dependent on macroH2A.

Discussion

DKK1 is involved in developmental signaling but is hardly expressed in adult tissues, so the expression of this gene in cancer cells is aberrant. DKK1 is upregulated in many cancer types and promotes malignant characteristics (Fatima, 2011; Kagey and He, 2017). In particular, high levels of DKK1 are found in tumors and serum of hepatocellular carcinoma patients and are associated with poor prognosis (Désert et al., 2016; Huang et al., 2014; Kim et al., 2015; Sato et al., 2010; Shen et al., 2012a; Tao et al., 2013; Yang et al., 2013; Zhang et al., 2014). Moreover, DKK1 promoted the migration and invasion capacity of HepG2 and other cancer cell lines (Chen et al., 2013; Huang et al., 2014; Tao et al., 2013). High DKK1 expression is observed in liver, breast and lung cancer cells with stem cell-like characteristics, suggesting that it could contribute to an undifferentiated phenotype (Chen et al., 2016; Malladi et al., 2016). Importantly, we detected increased levels of DKK1 in the extracellular media of cells lacking macroH2A, which demonstrates that the transcriptional changes are accompanied by an increased secretion of this protein. It will be interesting to test if the increased migratory capacity of HepG2 DKD cells depends on the upregulation of DKK1. In addition, we will study if low expression of macroH2A1.1 and macroH2A2 is associated with elevated DKK1 expression in other cancers where DKK1 is upregulated and macroH2A has been functionally implicated such as breast, colorectal, lung and melanoma (Corujo and Buschbeck, 2018; Kagey and He, 2017).

Interestingly, DKK1 has been implicated in an autocrine evasion mechanism used by cancer cells to avoid being cleared by natural killer cells (NK) (Malladi et al., 2016). However, NOD/SCID mice lack NK cell function. Nevertheless, this highlights the potential functions of secreted proteins in autocrine and paracrine signaling in the tumor microenvironment that go beyond tumor growth. It is clear that to study such effects and the potential role of macroH2A-regulated gene expression in modifying the tumor microenvironment would require the use of specialized *in vivo* or cell culture models.

11.3. MacroH2A can regulate the transcriptional response to hypoxia and inflammatory signaling

Treatment with the pro-inflammatory cytokine TNF α greatly upregulated the expression of DKK1 specifically in HepG2 cells lacking macroH2A. The overexpression of DKK1 dependent on macroH2A was much greater in xenografts than in HepG2 cells cultured in standard conditions. In addition, xenografts derived from DKD cells showed an upregulation in hypoxia-related genes and, to a lesser degree, inflammation-related

genes. We hypothesize that the microenvironment that the cells are exposed to when grown as xenografts in mice could induce hypoxic and inflammatory signaling that triggers the observed overexpression of DKK1. Notably, the immunosuppressed NOD/SCID mice used in our experiment lack adaptive immunity but retain some macrophage function, so it is possible that the xenografted cells were exposed to TNF α *in vivo* (Hu et al., 2011; Krepler et al., 2004; Shibata et al., 1998).

Solid tumors exhibit regions of low oxygen availability that trigger hypoxic transcriptional responses mainly mediated by the HIF (Hypoxia Inducible Factor) transcription factors (Bertout et al., 2008). Hypoxic stress can limit the growth of tumor cells but at the same time is associated with the acquisition of malignancy, specially increased invasiveness and metastatic potential (Finger and Giaccia, 2010). Among the pro-invasive secreted factors than can be induced by hypoxia are the lysil oxidase (LOX) and metalloproteinases (MMPs) (Finger and Giaccia, 2010; Gilkes et al., 2014). We have found LOX, MMP17, MMP20 and MMP24 upregulated in DKD xenografts ($\log_2FC > 1.5$). A recent study reports that macroH2A1.2 represses LOX expression in breast cancer cells by reducing their osteoclastogenic potential (Kim et al., 2018b). Hypoxia and inflammation are at interplay in physiological and pathological conditions (Cummins and Taylor, 2017). Indeed, the NF- κ B transcription factor is a major mediator of the transcriptional response to inflammatory stimulus such as TNF α , is activated by hypoxic conditions and regulates HIF (Bandarra and Rocha, 2013). The knock-down of macroH2A1 in HepG2 cells in another study was reported to induce features through the activation of the NF- κ B transcription factor that are compatible with a more stem cell-like phenotype (Lo Re et al., 2017).

In summary, we have analyzed in greater detail the transcriptional regulation of DKK1 by macroH2A in response to TNF α . However, transcriptomic data from xenograft samples suggest that macroH2A could have a more general role in the transcriptional response to hypoxic and inflammatory stimulus that is not limited to one gene. Inflammation and hypoxia are hallmarks of the tumor microenvironment that are involved in the malignant transformations of cancer cells and the activity of the immune system. We will study in the future the effect of removing macroH2A in HepG2 cells upon TNF α signaling and hypoxic conditions in a genome-wide transcriptomics approach.

11.4. The maintenance of epigenetic and transcriptional states by macroH2A might be mediated by the stabilization of 3D chromatin structures

The absence of macroH2A altered the three-dimensional interaction profile of the *DKK1* promoter and the activity of a distal regulatory element in response to an activating inflammatory stimulus. These changes are highly correlated with the increase in *DKK1* mRNA levels. So far, we have only determined the interactions of the *DKK1* promoter in standard cell culture conditions. It will be very relevant to study whether additional macroH2A-dependent changes occur in this locus upon treatment with TNF α to better understand the impact of macroH2A in the structural organization of the locus. Moreover, we would like to interrogate if macroH2A directly or indirectly affects the binding of transcription factors that mediate the response to TNF α , such as NF- κ B, or if it rather limits the scope of action of enhancers by maintaining certain interactions. For instance, TAD boundaries limit promoter-enhancer interactions to particular chromatin domains (Franke et al., 2016; Guo et al., 2015; Lupiáñez et al., 2015). In addition, transcription factor function can be facilitated or impaired by chromatin conformation, and the same time chromatin conformation can be altered by transcription factors (Stadhouders et al., 2019). HiC data in IMR90 cells treated with TNF α shows that the majority of responsive genes are already in close proximity to NF- κ B-targeted enhancers before the treatment, which suggests that the pre-existing chromatin structures are important for a proper transcriptional response (Jin et al., 2013).

Our results also show that macroH2A has a major function in constitutive heterochromatin architecture and promotes attachment to the nuclear lamina. Large heterochromatin domains in differentiated cells represent a major obstacle to reprogramming and de-differentiation events (Becker et al., 2016; Soufi et al., 2012; Wen et al., 2009). In addition, the association of chromatin with the lamina in LADs contributes to heterochromatin organization and genetic regulation during development and differentiation (Gonzalez-Sandoval et al., 2015; Peric-Hupkes et al., 2010; Robson et al., 2016). The stabilization of heterochromatin architecture and possibly other chromatin structures has the potential to provide a molecular explanation for the functions of macroH2A as an epigenetic stabilizer of differentiated cell states and as a barrier to reprogramming or transformation. We hypothesize that the transcriptional changes observed in the absence of macroH2A, in particular in the context of cancer, are in many cases secondary to the loss or modification of stable chromatin structures.

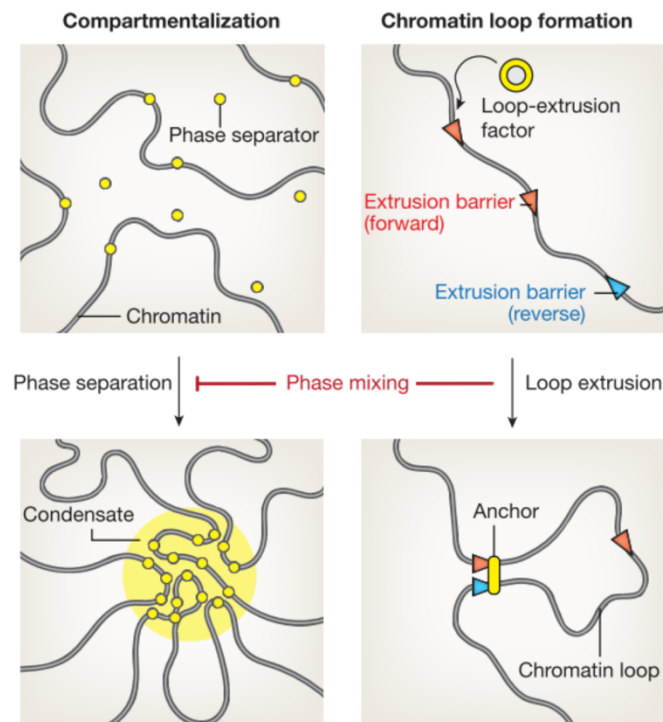
The role of macroH2A in heterochromatin organization and lamina attachment suggests that it might participate in the compartmentalization of broad domains rather than in the maintenance of specific chromatin loops. It has been recently proposed that phase separation could drive chromosomal compartmentalization and in particular the formation of heterochromatic domains (Larson and Narlikar, 2018; Larson et al., 2017; Strom et al., 2017). Phase separation describes a process in which proteins with intrinsically disordered regions self-organize into condensates that exhibit liquid-liquid separation. HP1, for instance, contains intrinsically disordered regions and has been shown to form oligomers that exhibit phase-separation properties both *in vivo* and *in vitro* (Larson et al., 2017; Strom et al., 2017). Interestingly, compartmentalization is envisioned as the “default” self-organization mechanism of the genome while loop-extrusion acts as an antagonizing mechanism that establishes insulated loops resistant to compartmentalization (**Figure 49A**) (Nuebler et al., 2018; Stadhouders et al., 2019). In support of this view, disruption of loop extrusion by cohesin or CTCF depletion leads to a loss of TAD organization and increased compartmentalization (Haarhuis et al., 2017; Nora et al., 2017). The linker region of macroH2A is an intrinsically disordered and highly charged region. It is speculative to think that the linker domain of macroH2A could drive a process similar to phase separation, but we suggest that it could participate in the compartmentalization of chromatin by promoting fiber-fiber interactions over large domains. The loss or destabilization of compartmentalization could lead to spurious loop extrusion events and result in aberrant contacts between genes and regulatory elements.

The epigenome and three-dimensional conformation of the genome can act as a barrier to phenotypic changes and stabilize cellular identities by maintaining robust transcriptional programs. It is suggested that a driving force of tumor progression and evolution is the alteration of epigenetic regulatory mechanisms and the loss of stable genomic organization, which leads to “epigenetic stochasticity” or “epigenetic plasticity” associated with increased “stemness” (**Figure 49B**) (Feinberg et al., 2016). This increased stochasticity results in epigenetic and transcriptional “noise” that promote tumor heterogeneity and the selection of malignant characteristics by the tumor environment. According to this model, this ultimately results in cancer progression and drug resistance. One study shows that macroH2A occupancy at transcription factor binding sites suppresses transcriptional noise in front of perturbations in the form of viral infections (Lavigne et al., 2015). We would like to suggest that the stabilization of chromatin

Discussion

structures by macroH2A could act as a mechanism to generally contribute to transcriptional robustness and explain its contribution to stable cell states.

A



B

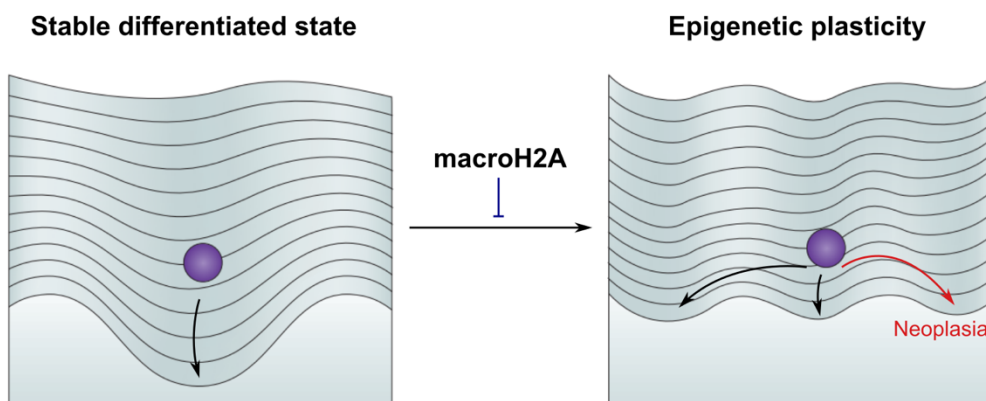


Figure 49. Model of epigenetic stability mediated by macroH2A. (A) Schematic representation of two major mechanisms of genome organization. In the left panels, a phase separator factor, which could be macroH2A, creates a condensate that exhibits phase separation properties with its surrounding environment. In the right panels, loop extrusion factors, like the cohesin complex, generate stable loops limited by boundary elements such as CTCF. The two mechanisms are thought to counteract each other. Panel taken from Stadhouders, 2019. **(B)** Illustration of the role of macroH2A in the stabilization of epigenetic states using Waddington's epigenetic landscape. Robust epigenetic regulatory mechanisms are necessary for a cell to achieve and maintain a stable differentiated state. Loss or alterations in epigenetic regulators increases stochasticity and “flattens” the barriers in the epigenetic landscape of the cell, which can lead to the loss of differentiated features or neoplastic transformations. We propose that macroH2A acts as a stabilizer factor in this model by maintaining chromatin organization. Panel adapted from Feinberg, 2016.

Conclusions

12. Conclusions

In the first part of this study, we studied the role of macroH2A in the nuclear organization of chromatin in HepG2 cells and came to the following conclusions:

1. MacroH2A removal results in a global loss of heterochromatin organization, increased frequency of nuclear defects and bigger nucleoli.
2. A substantial fraction of macroH2A is enriched at constitutive heterochromatin regions marked by H3K9me3 including repetitive elements. MacroH2A is not needed for the presence of H3K9me3 or the one-dimensional compaction of chromatin at these regions and participates in the transcriptional silencing of repetitive elements but is not the main factor.
3. MacroH2A is necessary for maintaining the organization of repetitive elements and their attachment to lamin B1 and promotes overall proximity of H3K9me3-marked chromatin to lamin B1.

In the second part of this study, we aimed to characterize the molecular mechanisms and domain requirements of the function of macroH2A in chromatin dynamics and organization. The conclusions we derive from the results are:

1. Conserved structural features allow macroH2A1.1, but not macroH2A1.2 or macroH2A2, to bind ADP-ribose and inhibit the enzymatic activity of PARP1.
2. All macroH2As can suppress DNA damage-induced chromatin expansion. MacroH2A1.1 has the strongest effect which reflects its specific capacity to inhibit PARP1, while the highly basic linker region of all macroH2As can limit chromatin expansion to a lesser degree in a PARP1-independent manner.
3. The macroH2A2 linker is essential and sufficient to maintain H3K9me3-marked heterochromatin architecture.

In the third and final part of my thesis, we wanted to describe the changes in the transcription and phenotype of cells in a cancer model that depend on the presence of macroH2A. The conclusions from this study are the following:

Conclusion

1. HepG2 cells lacking macroH2A expression have an increased colony formation and migratory capacity, but are unaffected in cell growth both in cell culture and when xenografted.
2. Transcriptomic profiling of tumors derived from HepG2 cells shows significant gene expression changes that depend on macroH2A, which includes deregulation of genes related to cellular adhesion, development, hypoxia and inflammation. This includes the upregulation of the cancer-promoting gene *DKK1*.
3. MacroH2A loss renders *DKK1* sensitive to activation in response to inflammatory TNF α signaling. The repression of *DKK1* can be rescued by macroH2A1.1 and macroH2A2, but not macroH2A1.2.

Materials and Methods

13. Materials and Methods

13.1. Antibodies

We generated specific antibodies against macroH2A1 as previously described (Buschbeck et al., 2009). Briefly, rabbits were immunized with bacterially expressed and purified His6- or glutathione S-transferase (GST)-tagged macro domains. After terminal bleeding, we affinity purified the antibodies using corresponding GST-tagged macro domains crosslinked to GSH-sepharose beads (GE Healthcare). All animal procedures were carried out by the CID-CSIC Antibody Generation Service. The complete list of home-made and commercial antibodies used is included in the antibodies table (13.17).

13.2. Plasmids

We used standard PCR and restriction-enzyme based cloning techniques to generate protein expression plasmids and shRNA expression plasmids. We used previously described pRetroSUPER constructs for shRNA-mediated knock-down of macroH2A1 and macroH2A2 (Buschbeck et al., 2009). The shRNA cassette directed against macroH2A2 was shuttled into a pRetroSUPER-hygromycin-resistant backbone to permit the generation of macroH2A1/2 double knock-down HepG2 cells. Two different lamin B1-specific shRNAs expressed from lentiviral PLKO.1-puro backbone were ordered from Sigma (TRCN0000297156, TRCN0000029270). The sequences of all shRNAs used can be found in **Table 1**.

Name	Sequence
sh mH2A1	GATCCCCCAGTTACTTCGTGTCTACTTCAAGAGAGTAGACACGAAGTAACTGGTTTTTGGAAAC
sh mH2A2	GATCCCCAACACAGCCGAAATTGACTTCAAGAGAGTCAATTTTCGGCTGTGGTTTTTTTGGAAAC
sh LMNB1 #1	CCGGGCTCAAAGAAGTACAGTCTTTCTCGAGAAAGACTGTACTTCTTTGAGCTTTTTTG
sh LMNB1 #2	CCGGCGCTTGAAGAACACTTCTGAACTCGAGTTCAGAAGTGTCTTCAAGCGTTTTT

Table 1. shRNA sequences

We generated mammalian expression constructs of mCherry-tagged human full-length macroH2A1.1, macroH2A1.2 and macroH2A2 by cloning the PCR amplified cDNA of each macroH2A histone into pmCherry-C3. Single residue mutants or whole domain deletions were generated using QuickChange site-directed mutagenesis (Stratagene). Human macroH2A2 constructs used to generate HepG2 cell lines were cloned into pBabe.puro plasmids (Morgenstern and Land, 1990) with an N-terminal YFP fusion tag.

Materials and Methods

The Δ LMD constructs contain amino acids 1-123 while Δ MD contain amino acids 1-161. Full-length FLAG-tagged mouse macroH2A1.1, macroH2A1.2 and macroH2A were generated by cloning the PCR amplified cDNA of each macroH2A histone into a pCMV-Tag2A and then subcloned into pBabe.puro. N-terminal His-tagged macrodomains of macroH2A1.1, macroH2A1.1 G224E and macroH2A2 were generated by cloning into pET28a. The histone H2A cDNA was provided by Sandra Hake and subcloned into pmCherry-C1. The histone H2B-PAGFP construct has been previously described (Beaudouin et al., 2006).

13.3. Cell culture

13.3.1. Standard cell culture conditions

Unless stated otherwise, the following conditions were used for cell culture. All cells were incubated at 37°C in 5% CO₂. HepG2 (ATCC, HB-8065), U2OS (ATCC, HTB-96), GP2-293 and HeLa cells (ATCC, CCL-2) were routinely cultured in Dulbecco's modified Eagle medium (DMEM) containing 4.5 g/L glucose (Gibco) supplemented with 10% v/v fetal bovine serum (FBS) (Gibco), 2 mM L-glutamine (Gibco), 50 U/ml penicillin (Gibco) and 50 mg/ml streptomycin (Gibco). MEF were cultured in DMEM containing 4.5 g/L glucose supplemented with 10% v/v FBS, 2 mM L-glutamine, 0.1 mM β -mercaptoethanol, 1x MEM non-essential amino acids solution (Gibco), 1mM Na-pyruvate (Gibco), 50 U/ml penicillin and 50 mg/ml streptomycin. mESC (E14Tg2A.4) were plated onto 0.1% gelatin-coated plates and cultured in minimal essential medium (MEM) (Sigma-Aldrich) supplemented with 20% FBS (HyClone), 0.1 mM β -mercaptoethanol, 1x MEM non-essential amino acids solution, 1mM Na-pyruvate, 50 U/ml penicillin, 50 mg/ml streptomycin and leukemia inhibitory factor (LIF) corresponding to 1000 U/mL. LIF was produced in house from a plasmid. HepG2, GP2-293 and mESC cells were periodically checked for mycoplasma contamination.

13.3.2. Colony formation assay

50 thousand HepG2 cells were seeded in six-well plates and grown in standard conditions for one week. The plates were then washed with PBS, fixed with paraformaldehyde (PFA) 4% in PBS for 10 min at room temperature (RT) and stained with crystal violet 0.01% in ethanol for 30 min at RT. Finally, the plates were destained by repeatedly submerging in abundant water, air-dried and scanned.

13.3.3. Wound healing assay

2 million HepG2 cells were seeded in six-well plates containing a silicone insert that creates a cell-free gap 500 μm wide (Ibidi). After two days the cultures reached confluency, the inserts were removed and the FBS in the media was reduced to 0.5% to limit cellular proliferation. The wounds were imaged at the same position at the time of insert removal (time 0h) and at 24h, 48h and 72h. Three wounds were generated in each experiment for each condition as technical replicates. The area of the wound at each timepoint was calculated with the ImageJ macro MRI Wound Healing Tool (http://dev.mri.cnrs.fr/projects/imagej-macros/wiki/Wound_Healing_Tool). For each wound, the wound area at each timepoint was divided by the area measured at time 0h.

13.3.4. XTT proliferation assay

The assay is based on the cleavage of the tetrazolium salt XTT in the presence of an electron-coupling reagent, producing a soluble formazan salt. This conversion only occurs in viable cells. 5 thousand cells were grown in a 96-well microplate and incubated with the XTT labeling mixture (50 μl /well) for 4 hours. After this incubation period, the formazan dye formed is quantitated using a scanning multi-well spectrophotometer, 492nm-650nm. The absorbance revealed directly correlates to the cell number.

13.3.5. $\text{TNF}\alpha$ treatment

500 thousand HepG2 cells were seeded in six-well plates and incubated overnight in standard conditions. Then, the cells were subject to an overnight serum starvation by reducing the FBS concentration to 0.5%. After that, $\text{TNF}\alpha$ (PeproTech, resuspended in PBS 0.1% BSA) was added to the media to a final concentration between 10 and 100 ng/mL (as indicated in each figure). An equal volume of PBS 0.1% BSA was added as the untreated control condition. The cells were treated for 24h, then collected and subject to protein and RNA expression analysis. The same protocol was followed to obtain samples for ChIP experiments but 10 million cells were seeded in P15 plates and grown for two days before applying the o/n serum starvation, subsequent $\text{TNF}\alpha$ treatment and fixation after 24h.

13.4. Gene transduction

13.4.1. Transfection

Materials and Methods

U2OS cells were transiently transfected with XtremeGENE HP DNA transfection reagent (Roche). The generation and culture of HeLa Kyoto cells with an inducible macroH2A1.1 transgene have been described previously (Timinszky et al., 2009). For induction, 1 µg/ml doxycycline (Clontech) was added 48 h before analysis.

GP2-293 cells were used as packaging cells to produce viral particles for retroviral infections. Four million GP2-293 cells were seeded in P10 plates and cultured to 60-70% confluency. At that point the cells were transfected with 10 µg of the retroviral plasmid of interest and 3 µg of the pCMV-VSV-G plasmid for pseudo-typing with VSV-G which allows the infection of a broad range of cells. Plasmidic DNA was mixed in a 1x HBS solution (2x HBS: 272 mM NaCl, 2.8 mM Na₂HPO₄, 55mM HEPES, pH 7) containing 125mM CaCl₂ in a total volume of 800 µL. The mixture was added onto the cell culture media dropwise and left o/n. Transfection efficiency was controlled by using a GFP-containing retroviral vector and the media was replaced by the target cell media. The culture media was collected at 24h, replaced by fresh media and collected again after 24h.

13.4.2. Retroviral infection

The supernatant containing viral particles produced by GP2-293 cells was filtered using a 0.45 µm filter and supplemented with 8 µg/mL of Polybrene (Sigma-Aldrich) to enhance the infection efficiency. Infection hits were performed by adding the filtered viral supernatant to the target cells cultured in six-well plates at 60-70% confluency. Cells were then centrifuged 45 min at 1200rpm at 37°C, incubated at 37°C for 45min and then cultured o/n in fresh media. The same process was repeated 24h after the first infection. The cells were selected with antibiotics depending on the resistance of the plasmid used (2 µg/ml puromycin, 150 µg/ml hygromycin and/or 30 µg/ml bleomycin). The necessary selection time was determined by using a negative control plasmid without resistance. This procedure was used to generate all HepG2 stable cell lines.

13.5. Protein expression analysis

13.5.1. Protein extraction from cultured cells

Cells were collected by trypsinization and washed twice with PBS. Cell pellets were directly resuspended with Laemmli's buffer and sonicated using a Bioruptor® Plus (Diagenode, 10 cycles of 30 sec ON/30 sec OFF at high intensity). Lastly they were incubated at 95°C for 5 min before western blot analysis.

13.5.2. Precipitation of extracellular proteins

10 million HepG2 cells were seeded and treated with TNF α as described in **13.3.5**. 1 mL of culture media was collected and used for acetone precipitation of extracellular proteins. First, the sample was briefly centrifuged to remove any detached cell that might be in suspension. Then, four volumes of cold (-20°C) acetone were added to the tube. After vortexing, the sample was incubated o/n at -20°C. The sample was centrifuged 10min at 13000g and the supernatant was discarded. The protein pellet was air-dried at RT for 30 min and resuspended in 100 μ L of Laemmli's buffer. After heating at 95°C for 10 min, the sample were analyzed by Western blot.

13.5.3. Protein extraction from mouse tissues

Fresh tissues from healthy 6 weeks old male NOD mice were kindly provided by Marta Vives, Arce Garcia-Jaraquemada and Josep Manyé. Organs were cleaned with PBS, frozen in liquid nitrogen and powdered using a liquid nitrogen-cooled mortar. In order to obtain total protein extracts, 0.1 g of powdered tissue was resuspended in 500 μ L Laemmli's buffer, sonicated with an ultrasonic homogenizer (Omni Ruptor 4000) and then boiled at 95°C for 10 min for western blot analysis.

13.5.4. Western blot (SDS-PAGE)

Samples were loaded on 12% polyacrylamide gels and run at 36 mA for 45-60 min to achieve optimal separation and then transferred at 220 mA for 90 min onto a nitrocellulose membrane (GE healthcare). After transfer, protein load was checked by Ponceau staining and the membranes were blocked using 5% non-fat milk (Nestlé) in TBST for 15 min. Membranes were incubated with primary antibodies o/n at 4°C. The day after, membranes were washed thrice with TBST for 10 min and incubated with appropriate secondary HRP-conjugated antibodies (DakoCytomation) for 1 h at RT. Membranes were then washed again thrice with TBST for 10min and incubated 1 min with Pierce™ ECL Western Blotting Substrate (Thermo Scientific) chemiluminescent reagent mix in a 1:1 ratio. Dried membranes were overlaid with photographic film (GE Healthcare) for a few seconds up to 5 min to achieve proper exposures and the films were developed with a FujiFilm FPM-100A developer.

As an alternative detection method, the same procedure was carried out but with the following modifications. Fluorophore-conjugated secondary antibodies (LI-COR Biosciences IRDye 680RD and IRDye 800CW) were used, which do not require any

Materials and Methods

substrate, and the dried membranes were scanned with an Odyssey® CLx Imager. The antibodies used for western blot are listed in the table at **13.17**.

13.6. RNA expression analysis

13.6.1. RNA isolation

Cultured cells were collected by trypsinization and washed twice with PBS. Total RNA was isolated using either the PureLink™ RNA Mini kit (ThermoFisher Scientific) or the Maxwell® RSC simplyRNA Cells Kit (Promega) together with the Maxwell® RSC Instrument (Promega) according to the manufacturer's instructions. Total RNA from xenograft samples was isolated using the TRIzol™ Reagent (ThermoFisher Scientific). Concentration and quality of the extracted RNA was checked using a Nanodrop instrument (Thermo Scientific).

13.6.2. cDNA synthesis by reverse transcription

1 µg of total RNA was used for cDNA synthesis using the First Strand cDNA synthesis kit (Thermo Scientific) according to the manufacturer's instructions. In most samples oligo(dT) primers (1 µg) were used to synthesize cDNA from poly(A) RNA. For the expression analysis of repetitive elements, random hexamer primers (1 µg) were used and the mixture of RNA with primers was incubated 5 min at 65°C before the addition of the reverse transcriptase and the rest of reaction components. All reverse transcription reactions were incubated for 1h at 37°C and terminated by heating at 72°C for 5 min.

13.6.3. Quantitative real-time PCR (RT-qPCR)

cDNAs were diluted 10 fold previous to real-time PCR analysis using a LightCycler® 480 II instrument (Roche). A 10 µL reaction mix containing 5 µL LightCycler® 480 SYBR Green I Master, 0.5 µL of forward and reverse primers (for a final concentration of 0.5 µM), 2 µL of diluted cDNA and 2 µL of autoclaved ddH₂O was loaded in 384-well plates. All samples were analyzed in technical duplicates. Autoclaved ddH₂O was used as a negative control. The qPCR reaction conditions are summarized in **Table 2**.

The median efficiency of each primer pair was determined in every RT-qPCR plate run using the online software chainy (Mallona et al., 2017) which is mainly based on the R packages qpcR and NormqPCR (Perkins et al., 2012; Ritz and Spiess, 2008). The expression of a target RNA in a sample relative to a control sample was calculated using a variation of the ΔC_t method that accounts for the efficiency of the primers as follows:

$$efficiency_{target}^{(Ct_{control}-Ct_{sample})}$$

All values were normalized by the average expression of two reference genes (“housekeeping” genes): *GAPDH* and *RPLP0*. The normalized relative expression value is used in all the plots and statistics. All quantifications of RNA expression are relative, so that the chosen control sample will have an expression of 1 for all target genes and the rest of the samples will have an expression that represents a multiple or fraction of the expression in the control.

Step	Temperature (°C)	Duration	Cycles
Preincubation	95	8 min	1
Amplification	95	10 sec	45
	62	20 sec	
	72	15 sec	
Melting curve	95	15 sec	1
	62	1 min	
	99		
Cooling	40	30 sec	1

Table 2. Conditions for qPCR reactions.

13.6.4. *In situ* mRNA hybridization

Tissues were fixed in 10% formalin. Hybridization with 35S-labeled antisense RNA probes was carried out on 20 µm cryosections, as described previously (Treier et al., 1998). Hybridization signals were detected by autoradiography using Kodak NTB-2 liquid emulsion. Exposure time varied from 2 to 4 weeks depending on signal strength, as determined by previous exposure to Kodak BioMax MR films. Tissue sections were counterstained with 0.001% bisbenzimidazole in PBS. Pictures of signal and counter stain were overlaid using Adobe Photoshop software.

13.6.5. RNA-Seq and differential gene expression analysis

1 µg of total RNA extracted from xenograft tumors was submitted to the Genomics Unit of the Centre for Genomic Regulation (Barcelona, Spain) who prepared the libraries

Materials and Methods

with a poly(A) selection and sequenced them in an Illumina HiSeq2500 using 50bp single reads. Five different samples of each X.RDM and X.DKD were sequenced.

To quantify the transcript expression we used Salmon version 0.14.0 (Patro et al., 2017), using libtype A and “validateMappigs” option active. The raw reads were “quasi-aligned” to the human Ensembl genome annotation GRCh38.12. The quantification results obtained from Salmon were imported to the DESeq2 R package (Love et al., 2014) using the tximport R package (Soneson et al., 2016). The design formula for DESeq2 included only the type of tumor as a variable (X.RDM or X.DKD) and the contrast X.DKD vs X.RDM was applied to obtain log2FC and p-values. Genes with a Benjamini-Hochberg adjusted p-value lower than 0.05 were selected as being significantly differentially expressed. In some cases an additional log2FC cut-off was used as specified in the analysis. Genes were annotated with the biomaRt R package. PCA, M-A and volcano plots were generated using ggplot2 (Wickham, 2016).

13.6.6. Gene ontology enrichment analysis

Significant differentially expressed genes with a $\log_2FC \geq 1$ and a HGNC annotation were selected for a gene ontology enrichment analysis. The function getEnrichedGO from the R package ChIPpeakAnno was used, with a p-value cut-off of 0.05 and minGOTerm = 10 as parameters (minimum occurrences of a GO term in the genome for it to be included in the analysis).

13.6.7. Gene Set Enrichment Analysis (GSEA)

All differentially expressed genes with a HGNC annotation were ranked by log2FC value for a Gene Set Enrichment Analysis (GSEA) using the pre-ranked functionality of the GSEA software (v4.0.1 for Mac) (Mootha et al., 2003; Subramanian et al., 2005). As gene database we selected the hallmark gene sets from the Molecular Signatures Database (MSigDB v6.2 updated July 2018) (Liberzon et al., 2011, 2015) which contains gene expression signatures derived by aggregating many MSigDB gene sets to represent well-defined biological states or processes.

13.7. Chromatin immunoprecipitation (ChIP)

13.7.1. ChIP-qPCR

Between 20 and 30 million HepG2 cells were trypsinized and fixed in a PBS solution containing 1% PFA and 10% FBS for 10 min at RT in rotation. The fixation was

quenched by the addition of glycine to a concentration of 0.125M and rotation for 5 min at RT. Cells were then washed with cold PBS and either lysed directly or stored at -80°C. Nuclei extraction was achieved by resuspending the fixed cells in 500 µL lysis buffer I (5 mM PIPES pH 8, 85 mM KCl, 0.5% NP-40, 1 mM PMSF, 50 µg/mL leupeptin) and incubating for 30min on ice with occasional gentle pipetting. Lysed cells were centrifuged at 5000 rpm for 5 min at 4°C, the supernatant was discarded and the pellet resuspended in 500 µL lysis buffer II (1% SDS, 10 mM EDTA pH 8, 50 mM TRIS pH 8, 1 mM PMSF, 50 µg/mL leupeptin) and incubated for 30min at 4°C in rotation. Chromatin shearing was achieved by sonication in a Bioruptor® Plus (Diagenode, two rounds of 10 cycles 30 sec ON/30 sec OFF at high intensity) or a Bioruptor® Pico (Diagenode, 10 cycles 30sec ON/30 sec OFF at high intensity). After sonication, the samples were centrifuged for 10 min at 10000rpm to remove insoluble debris. From this point on all samples were kept on ice until specified otherwise. A 10 µL aliquot of lysate was mixed with 90 µL of water, de-crosslinked for 1 h at 65°C, purified using a PureLink™ Quick Gel Extraction and PCR Purification Combo Kit (Invitrogen), quantified using a Nanodrop (Thermo Scientific) and run in a 1% agarose gel electrophoresis. This allowed to check the size of the chromatin fragments (ranging 300-500bp) and estimate the concentration of DNA in the lysate as a proxy for chromatin amount. 35 µg of sheared chromatin were used for each immunoprecipitation (IP) reaction diluted with 9 volumes of IP dilution buffer (1% Triton X-100, 150 mM NaCl, 2 mM EDTA pH 8, 20 mM TRIS pH 8, 1mM PMSF).

Samples were precleared with 20 µL of a mix of proteins A and G agarose beads containing salmon sonicated sperm DNA (Merck Millipore) or 20 µL of Magna ChIP™ protein A/G magnetic beads (Merck Millipore) for 2 h at 4°C in rotation. 10% of the precleared lysate was taken as input and stored on ice. Between 1 and 5 µg of antibody was added and the IP samples were incubated o/n at 4°C in rotation. The following day 20 µL of a mix of proteins A and G agarose beads containing salmon sonicated sperm DNA (Merck Millipore) or 20 µL of Magna ChIP™ protein A/G magnetic beads (Merck Millipore) were added to capture the antibody-bound chromatin fraction for 2 h at 4°C in rotation. Beads were spun down (3 min at 300g) or separated with a magnet (DynaMag™-2 Magnet) and washed at RT with each of the following buffers in this order: mixed micelle was buffer (140 mM NaCl, 20 mM Tris-HCl pH 8, 5 mM EDTA pH 8, 5% sucrose w/v, 1% Triton X-100 v/v, 0.2% SDS, 0.02% NaN₃), LiCl/detergent wash (0.5% deoxycholic acid sodium salt w/v, 1 mM EDTA, 250 mM LiCl, 0.5% NP-40 v/v, 10 mM Tris-HCl pH 8) and buffer 500 (0.1% deoxycholic acid sodium salt w/v, 1mM EDTA, 50mM HEPES pH 7.5,

Materials and Methods

500 mM NaCl, 1% Triton X-100 v/v, 0.02% NaN₃). Two washes were performed with each buffer, first just quickly resuspending the beads and the second in a rotating wheel for 3 min. Finally, a quick wash with TE buffer (10mM Tris-HCl pH8, 1 mM EDTA pH 8) was performed. The enriched chromatin was eluted by resuspending the beads in 200 µL of elution buffer (1%SDS, 100 mM NaHCO₃), 30 sec of vortexing and incubation for 30 min at RT in rotation. Two sequential elutions were performed and the eluates were combined. The input samples were included again in the experimental procedure at this point. Crosslinking reversal was achieved by adding NaCl to a final concentration of 200 mM and incubating the samples o/n at 65°C shaking. The following day the protein in the samples was digested by adding 2 µL of Proteinase K 10 mg/mL, 16 µL 1M TRIS pH 6.5 and 8 µL 500 mM EDTA and incubating for 2 h at 45°C.

After digestion, the DNA was purified by a phenol/chloroform extraction followed by ethanol precipitation. Shortly, one volume of phenol/chloroform/isoamyl alcohol (25:24:1, Sigma-Aldrich) was added to the sample and after vigorous shaking and centrifugation the upper aqueous phase was recovered. This was repeated two times and followed by addition of one volume of chloroform (Sigma-Aldrich). After shaking and spinning again, the aqueous phase was recovered and the DNA was precipitated by adding 0.05 volumes of 5M NaCl, 1 µL of glycogen (Roche) and 3 volumes of ethanol. After incubation at -80°C for 30min up to o/n, DNA was pelleted by centrifugation at >10000g for 30min, washed with cold 70% ethanol and air-dried. The DNA pellet was finally resuspended in 60 µL of water and analyzed by qPCR in a similar way that the RNA expression procedure (13.6.3) with slight modifications. The input samples were diluted 10 fold before loading into the plate. Efficiency of each primer pair was determined and used to calculate the relative ChIP enrichment. The input of each sample was used as the “control” to obtain a relative enrichment over input, using the same formula than for RNA expression. The relative enrichment was then converted to a percentage. The percentage of input value is used in all the ChIP plots, in some cases divided by the signal of the H3 ChIP if specified.

13.7.2. ChIP-Seq and association analysis

We followed the same procedure described for ChIP-qPCR but four IP reactions were pooled after DNA precipitation to achieve a greater amount of immunoprecipitated DNA suitable for ChIP-Seq. DNA enriched by ChIP was fluorometrically quantified with PicoGreen (Thermo Fisher) and 10 ng were used for library preparation. Library

generation and direct massive parallel sequencing on an Illumina genome analyzer were performed at the core facility of the EMBL (Heidelberg, Germany). The reads obtained were cleaned based on quality, and trimmed using the software tool Sickle (<https://github.com/najoshi/sickle>) and Cutadapt (Martin, 2011), and aligned versus the human genome (GRCh37/hg19) using Bowtie 2 version 2.0.6 (Langmead and Salzberg, 2012), with the sensitive preset option (-D 15 -R 2 -L 22 -i S,1,1.15). To detect genomic regions enriched for multiple overlapping peaks, MACS software version 1.4.2 was used (Zhang et al., 2008) with a p-value cutoff of 1.0×10^{-5} and a false discovery rate (FDR) of 5%. The enrichment profiles were calculated using the ngsplot ver 2.0 program (Shen et al., 2014). Screenshots of regions of interest were taken from the UCSC genome browser (Casper et al., 2018). To calculate the association between regions (peaks or DNA elements) we used regioneR, an R package based on permutation tests (Gel et al., 2016). To compare the associations between them, the calculated z-scores were normalized by dividing them by \sqrt{n} , where n refers to the number of regions in the region sets tested. We have further analyzed the following data sets generated by the ENCODE Project Consortium (ENCODE Project Consortium et al., 2013): HepG 2×75 Sg 2 RNA-seq (GEO, GSM958732; UCSC, wgEncodeEH000127); HepG2 H3K9m3 (GEO, GSM1003519; UCSC, wgEncodeEH003087); HepG2 H3K27me3 (GEO, GSM733754; UCSC, wgEncodeEH001023); and Repeatmasker (Smit, A., Hubley, R. and Green, P.; RepeatMasker Open-4.0. 2013-2015; <http://www.repeatmasker.org/>). ChIP-seq data has been deposited in the Gene Expression Omnibus (GEO) under accession number GSE58175.

13.8. UMI-4C

2.5 million HepG2 cells were seeded in 6-well plates. The day after, the cells were fixed by addition of DMEM 10% FBS 4% PFA and incubation for 10 min at RT. The reaction was stopped by adding glycine to a final concentration of 0.125M and incubation for 5 min on ice. The cells were then collected by trypsinization, washed with PBS and stored at -80°C. Later, cells were thawed, resuspended in 500 µL of lysis buffer (10 mM TRIS pH 8, 10 mM NaCl, 0.2% NP-40) and incubated for 30 min on ice. Nuclei were isolated by centrifugation for 5 min at 760 g and resuspended in 200 µL of DpnII buffer (NEB). 6 µL of SDS 10% were added to the tube and incubated for 60min at 37°C. Then, 20 µL of Triton X-100 were added and the sample was incubated for 60 min at 37°C. At this point, a 5 µL sample aliquot was collected as undigested control. Afterwards, digestion

Materials and Methods

of the nuclei was achieved by addition of 500U DpnII (NEB) and o/n incubation at 37°C. The next morning, 500U DpnII were added again and the sample was incubated for 4 more hours at 37°C to ensure digestion was as complete as possible. A 5 µL sample was taken at this point as the digested control and DpnII was inactivated for 15min at 65°C. Both undigested and digested controls were purified by phenol chloroform extraction and run in an agarose gel to check the correct digestion of the genomic material. Digested nuclei were pelleted by centrifugation for 5 min at 600 g and resuspended in 200 µL T4 ligation buffer (NEB) containing 20 mg/mL BSA (NEB) and 3 µL T4 ligase (NEB, 400 000 U/mL). The samples were incubated for 4 h at 16°C, then 5 µL T4 ligase were added and an overnight incubation at 16°C was carried out. The day after, a 5 µL aliquot was purified by phenol chloroform and run in an agarose gel to check the ligation of the previously digested fragments had occurred. The samples were then digested with 5 µL RNase A (10 mg/µL) and 10 µL proteinase K (10 mg/mL) for 30min at 37°C and decrosslinked o/n at 65°C. DNA was purified with AMPure XP beads (Beckman Coulter) and quantified with the Qubit™ dsDNA HS Assay Kit (Thermo Fisher). A total of 5 µg of DNA were sonicated with a Covaris S2 instrument with the following parameters: duty cycle 10%, intensity 4, cycles per burst 200 and time 60 s. Sequencing libraries were prepared as described in (Schwartzman et al., 2016). The procedure is based in a nested PCR using oligonucleotides aligned to the *DKK1* gene promoter which effectively functions as the “bait” or “point of view” for capturing contacts. The sequences of the primers used can be found in the primers table (13.18). Libraries were sequenced in an Illumina MiSeq instrument.

Sequencing data of the UMI-4C experiment was analyzed using the umi4cpackage (<https://bitbucket.org/tanaylab/umi4cpackage/src/default/>), an R package developed by the group of Amos Tanay as a pipeline specific for UMI-4C data (Schwartzman et al., 2016). Sequencing data from two experimental replicates was combined and analyzed with the default parameters of the package.

13.9. Microscopy

13.9.1. Immunohistochemistry

Immunohistochemical stainings of successive slides of tissue microarrays shown in were performed as described in Cantariño et al., 2016. The immunohistochemical analysis of HepG2-derived xenograft samples were performed by the Anatomical Pathology Service of the Hospital Germans Trias i Pujol according to standard protocols.

13.9.2. Immunofluorescence staining

For immunofluorescence staining, cells were grown on polylysine-coated slides (Menzel) and fixed in PBS with 4% paraformaldehyde (PFA) for 10 min at RT, permeabilized in ice-cold 0.1 M HCl with 0.5% Triton X-100 for 10 min at RT and blocked with PBS with 0.1% Tween (PBT) and 5% bovine serum albumin (BSA) for 1 h at RT. Primary antibodies were diluted in PBT with 5% BSA and incubated for 1 h at RT. After successive washes with PBT, secondary antibodies conjugated to Alexa Fluor 488 or 594 (Thermo Fisher) were diluted in PBT with 5% BSA and incubated for 30 min. After successive washes with PBT and PBS, slides were mounted with Vectashield® Mounting Medium with DAPI.

13.9.3. Fluorescence In Situ Hybridization (FISH) and immuno-FISH

Fluorescence in situ hybridizations (FISH) and slide preparations were performed essentially as described in Bolland et al., 2013, which describes a protocol to preserve as much as possible the three-dimensional morphology of the nucleus during the FISH experiment. Briefly, HepG2 cells were grown on polylysine-coated slides (Menzel) and fixed with PBS 4% PFA for 10 min at RT. The fixation was quenched by incubation with 0.1 M Tris-HCl pH 7.4 for 10 min at RT. After washing with PBS, the slides were incubated for 20 min in PBS 20% glycerol at RT. The slides were then subject to three cycles of freeze-thaw in liquid nitrogen, washed with PBS and incubated with 0.1 M HCl for 30 min at RT. For immuno-FISH, the immunofluorescence protocol was performed at this stage as described in **13.9.2**, starting with the blocking and finishing with the washes after secondary antibody incubation. Then, cells were washed with PBS and permeabilized in PBS 0.5 % saponin 0.5% Triton X-100 for 30 min at RT. Afterwards, slides were washed twice with PBS and incubated with 2x SSC 50% formamide for 10 min at RT (SSC 1x contains 150mM NaCl and 15 mM sodium citrate). At this point 20 ng of FISH probe in 20 μ L of hybridization solution (Sigma-Aldrich) was used for hybridization. The slides were heated to 78°C for 2 min and incubated o/n at 37°C in a humidified chamber. The day after, the slides were washed sequentially as follows: 15 min at 45°C in 2xSSC 50% formamide, 15 min at 63°C in 0.2xSSC, 5 min at 45°C in 2xSSC, 5 min at RT in 2xSSC and 5 min at RT with PBS. Finally, slides were mounted with Vectashield® Mounting Medium with DAPI.

FISH probes were PCR-labeled using dUTP-biotin (Roche) and detected with streptavidin conjugated to Alexa Fluor 488 (Thermo Fisher) or Tyramide Signal

Materials and Methods

Amplification (Thermo Fisher). The primers used to generate the FISH probes are listed in the primers table (13.18).

13.9.4. Proximity ligation assay (PLA)

Cells were grown on 12 mm coverslips. Prior to fixation in PBS with 4% PFA for 10 min at RT, cytoplasm was removed with CSK buffer (100 mM NaCl, 300 mM sucrose, 3 mM MgCl₂, 10 mM PIPES, 1:200 leupeptin and 1:100 PMSF) for 5 min on ice. Nuclei were permeabilized in 0.1 M HCl with 0.5% Triton X-100 for 10 min at RT and blocked with PBT with 5% BSA for 30 h at RT. Primary antibodies were diluted in PBT with 5% BSA and incubated overnight at 4°C. Then, detection of protein interactions was performed using the Duolink® PLA system (Sigma-Aldrich) according to the manufacturer's instructions.

13.9.5. Fluorescence recovery after photobleaching (FRAP)

FRAP experiments were performed with an inverted Leica SP8X WLL microscope equipped with a WLL2 laser (470 - 670 nm) and acusto-optical beam splitter. Live cells were recorded at 37°C in supplemented Leibovit's L15 medium. Images were acquired with a 40x1.3 objective at 600Hz. Cells were bleached for 10 frames using four different wavelengths (585 nm, 577 nm, 569 nm and 561 nm).

13.9.6. Transmission electron microscopy

Cells were washed in PBS (pH 7.2) and fixed in 1.6% glutaraldehyde diluted in 0.1 M PBS (pH 7.2) at room temperature; after 1 h, they were washed in the same buffer. The samples were acetylated according to (Thiry et al., 1985) before being embedded in Epon. Ultrathin sections mounted on copper grids were stained with uranyl acetate and lead citrate before examination with a Jeol JEM 1400 transmission electron microscope (TEM) at 80 kV.

13.9.7. Imaging of immunofluorescence staining and FISH

Confocal optical z-stacks of images were obtained by using a Zeiss LSM 710 confocal laser-scanning microscope equipped with a 63x1.4 NA Plan- Apochromat oil immersion objective. Fluorochromes were excited with the following laser lines: 405 nm (for DAPI detection), 488 nm (for Alexa Fluor 488 detection) and 594 nm (for Alexa Fluor 594 detection). Laser intensity and gain parameters were set to ensure proper exposure without saturated pixels. Ribosomal DNA FISH images were obtained using a Zeiss

AxioObserver Z1 fluorescence microscope equipped with a 63×1.4 NA Plan-Apochromat oil immersion objective. All acquisition parameters were kept equal during imaging of each experiment.

13.9.8. Imaging of DNA damage induced by laser micro-irradiation

U2OS cells were plated on Lab-Tek chambered coverglass (Nunc). For the induction and imaging of DNA damage, cells were pre-sensitized by addition of 0.3 µg/mL Hoechst 33285 into the culture media for 1 h. When indicated cells were treated with 1 µM Olaparib (Selleckchem) for 1 h. Immediately before imaging, the growth medium was replaced by Leibovit's L-15 medium, supplemented with 10% fetal bovine serum, 2mM glutamine, 50U/ml penicillin and 50mg/ml streptomycin. Laser microirradiation experiments were performed on a Zeiss AxioObserver Z1 confocal spinning-disk microscope equipped with an AxioCam HRm CCD camera (Zeiss) or a sCMOS ORCA Flash 4.0 camera (Hamamatsu) through a Zeiss Plan/Apo 63Å~1.4 water-immersion objective lens. Laser micro-irradiation and PA-GFP activation were performed simultaneously using 405nm laser set to 100µW for 350 milliseconds, subsequently the images were taken every 4 sec. using 488nm laser, for 3 minutes.

13.9.9. Image analysis

Changes in chromatin expansion were analyzed automatically using a custom-made routine written in MatLab (MathWorks) as previously described in Sellou et al., 2016. In particular, we have measured the width of the microirradiated chromatin as the minor axis of an ellipsoid fitted to the area of the fluorescent signal of PAGFP-H2B. The measurements at 120s are normalized by the 1s timepoint to achieve the final "chromatin expansion" measure.

HepG2 immunofluorescence and FISH images were loaded and analyzed in Fiji, a distribution of ImageJ (Schindelin et al., 2012), using custom macro scripts. Image stacks were converted to maximum intensity z-projections and the DAPI signal was automatically thresholded with built-in methods and used to define and measure individual nuclei with particle analyzer. Similarly, FISH signals were automatically thresholded with built-in methods. For the analysis of the number of H3K9me3 foci (**Figure 36**), the H3K9me3 channel signal was cropped for each individual nucleus and analyzed by subtracting the background with a rolling-ball algorithm, applying a median filter and automatically thresholding the signal. Artifacts were removed after applying thresholding with noise

Materials and Methods

removal and binary processing functions of the software. FISH signals and H3K9me3 foci were automatically identified and counted or measured with the particle analyzer function of Fiji. Nucleoli were manually segmented from maximum intensity-projections of NPM1 immunostaining images.

For the PLA experiment, 3D stack-images were acquired by confocal microscopy. Then, cell nuclei were segmented from the DAPI channel z-stacks by using an automatic threshold (default built-in method) and a watershed algorithm. PLA signal blobs were identified from the green channel image stacks and assigned to the corresponding segmented nuclei with the FindFoci plugin (Herbert et al., 2014) with optimized parameters.

All images of the different experimental conditions were analyzed under the same parameters. Single cell measurements were grouped and plotted as box plots using either a web-tool developed by Tyers and Rappsilber laboratories (<http://boxplot.tyerslab.com>) or R base graphics and ggplot2 (Wickham, 2016).

13.10. Formaldehyde-assisted isolation of regulatory elements

Formaldehyde-assisted isolation of regulatory elements (FAIRE) assay was performed essentially as described in Simon et al., 2012. Briefly, sheared chromatin was prepared in the same way than for ChIP (13.7.1), 10% of the sample was kept as input and the rest was subject to two rounds of phenol/chloroform extraction, crosslink reversal and ethanol precipitation. Regions of interest were analyzed by qRT-PCR, the primers used are indicated in the primers table (13.18).

13.11. Protein expression and purification

For crystallographic analysis, the macrodomain of human macroH2A2 (aa 177-372) were cloned into a modified pET24 (Novagen), providing a tobacco etch virus (TEV) cleavable N-terminal 6xHis tag. The protein was expressed in *E. coli* BL21 (DE3) cells at 37 °C for 3h after induction with 1 mM isopropyl- β -D-thiogalactopyranoside (IPTG). Cells were harvested by centrifugation, resuspended in lysis buffer (50 mM NaH₂PO₄, 500 mM NaCl, 1 mM MgCl₂, 5mM β -mercaptoethanol, 10 % (w/v) glycerol), lysed with Emulsiflex-C5 (Avestin, Canada), and purified by Ni²⁺ sepharose (GE Lifesciences) affinity chromatography using the batch method. Beads were then washed in lysis buffer containing 1 M NaCl and 40 mM imidazole, followed by step-elution with lysis buffer

supplemented with 500 mM imidazole. Elution fractions were dialyzed against buffer containing 50 mM Tris pH 8.0, 300 mM NaCl, and in the presence of 6xHis-tagged TEV protease (1:100 molar ratio). The 6xHis tag, TEV protease and remaining impurities were removed by re-adsorption of the dialyzed protein to Ni²⁺ sepharose resin, followed by size exclusion chromatography on a Superdex 75 HR26/60 column (GE Lifesciences), equilibrated in 50 mM Tris pH 8.0, 300 mM NaCl. Monomeric peak fractions were dialyzed against 20 mM Bis-Tris pH 7.0, 100 mM NaCl, concentrated to 10 mg/mL (macroH2A2) and immediately used for crystallization.

For the zebrafish macrodomain construct, we carried out sequence alignments between fish and human macroH2A histones and chose domain boundaries identical to those of human macroH2A constructs. PCR fragments were cloned into a GST-fusion plasmid and the proteins expressed in *E. coli* at 18 °C, and purified using glutathione-coupled sepharose beads and lysis/wash buffer containing 0.5 M NaCl (50 mM Tris, pH 7.9, 0.1 mM MgCl₂, 1 mM EDTA and 1 mM DTT). Purified macrodomains were cleaved from GST using TEV protease.

For the production of human macroH2A1.1, macroH2A G224E and macroH2A2 macrodomains, proteins were produced in *E. coli* BL21 (DE3). Protein production was induced with 500 µM IPTG and bacterial culture grown overnight at 20°C. The next day, the bacteria were pelleted by centrifugation at 20 000 g for 30 min at 4°C. The bacterial pellet was lysed with lysis/wash buffer (50mM Tris-HCl, 300mM NaCl, 10mM imidazole, pH 8, protease inhibitors and lysozyme) and sonicated with an Omnisruptor 4000 equipped with a 3/4" tip (Omni International, 30"ON/ 30"OFF, 10 cycles). Once sonicated, the lysate was incubated with Ni-NTA beads (Qiagen) for 2 hs. After being washed with lysis buffer with 10 mM imidazole, proteins were eluted with elution buffer (50 mM Tris-HCl, 300 mM NaCl, and 200 mM imidazole). Proteins were dialyzed o/n against dialysis buffer (PBS 10% glycerol). The day after, protein concentration was checked by Nanodrop (Thermo Scientific) and stored at -80°C.

13.12. Protein crystallization and structure determination

MacroH2A2 macrodomains crystallized in sitting drop setups using 12.5% polyethylene glycol 1,000, 0.2 M Na(CH₃COO), 0.1 M MES (pH 6.5) as crystallization buffer. Ethyleneglycol was added to a final concentration of 15 % (v/v) and crystals of up to 200x40x40 µm were flash-frozen in liquid nitrogen. A dataset of 1.65 Å resolution has been recorded at beam-line PX01, Swiss Light Source, Villigen, Switzerland. The

Materials and Methods

structure of the macroH2A1.1 macrodomain has been previously determined (Kustatscher et al., 2005) and is deposited in PDB with the assigned ID 1ZR3. Briefly, crystals of an engineered form of macroH2A1.1 omitting the last ten C-terminal residues developed in 20% (w/v) polyethylene glycol, 0.2 M KCl in sitting drops using micro-seeding protocols. Platre-like crystals of up to 500 x 300 x 30 μm were transferred in reservoir solution supplemented with 20% v/v) glycerol and directly frozen in the cryo-stream at ESRF beam line ID29. Data processing and scaling were carried out with XDS (Kabsch, 1993).

The molecular replacement method as implemented in PHASER (Mccoy, 2007) was used to solve the structure of macroH2A in all crystal forms using the previously determined human macrodomain as search model (Kustatscher et al., 2005), PDB-ID 1ZR3). The structures were completed in alternating cycles of manual model correction in COOT and restrained TLS-refinement in autoBUSTER (Global Phasing Limited). The relatively high free R-factor observed with monoclinic macroH2A2 macrodomain crystals is likely due to one of two macrodomains in the asymmetric unit being partly disordered. Inspection of the refined model with MolProbity revealed excellent stereochemistry. Structural visualization was done with POVSCRIPT and POVRAY (<http://www.povray.org>). The protein structure data for the macroH2A2 macrodomains has been deposited to the protein data bank database [<https://www.rcsb.org>] and assigned the identifier 6FY5.

13.13. Isothermal titration calorimetry (ITC)

Binding of nucleotides to all macrodomain constructs was assessed using isothermal titration calorimetry (ITC) as previously described in Kustatscher et al., 2005. Assays were conducted using a VP-ITC instrument (MicroCal) at 25 °C. Proteins were dialyzed into ITC buffer (50 mM KH_2PO_4 , pH 6.5, 1 mM DTT). ADP-ribose was purchased from Sigma. MacroH2A macrodomains were at 45-60 μM concentration, whereas ADP-ribose was at 405-800 μM . Experimental data-fitting was performed using the Origin software (Originlab, USA).

13.14. PARP1 *in vitro* activity assay

The PARP1 auto-PARylation assay was performed in a buffer containing 50mM Tris-HCl (pH 8), 50 mM NaCl and 1 mM MgCl. All reactions contained 0.2 units/ μL PARP-1 HSA enzyme (Trevigen), 0.3x activated DNA (from 10x activated DNA, Trevigen) and 200 μM NAD^+ , but specific amounts of purified macro domains. Reactions were incubated

for 20min and stopped by addition of Laemmli's sample buffer and boiling at 95°C for 10 min. Samples were then separated on SDS-PAGE and analyzed by immunoblotting.

13.15. Xenograft

The animal experimentation procedure used for the xenograft experiment has been approved by the Animal Experimentation Commission of the Catalan Government (Comissió d'Experimentació Animal de la Generalitat de Catalunya) with project reference 9500. Four weeks old NOD/SCID immunosuppressed mice were purchased from Charles Rivers Laboratories (strain code 394). The animals were kept in specific-pathogen-free (SPF) conditions for the duration of the experiment. The animals were kept in the stall for two weeks before the experimental procedure. RDM and DKD HepG2 cells were resuspended in a Matrigel (Corning) and PBS mix 1:1. The animals were anesthetized with isoflurane and two subcutaneous injections were performed in each lower flank of the animal consisting of 100 μ L Matrigel:PBS 1:1 containing 5 million cells. A total of 5 mice (three males, two females) per condition was used, effectively resulting in 10 injections for each cell line. The tumor size was monitored every 2-3 days and measured with a digital caliper. Two measurements were performed for each tumor approximating the longer (a) and shorter (b) axis of an ellipse fitted to the tumor and the tumor volume was estimated by calculating $(a \times b)^2/2$. The weight and overall condition of the mice were regularly monitored. 21 days after the injection, tumors were resected with the mice under anesthesia and all animals were sacrificed. Tumors were cleaned with PBS and cut in half. One half was fixed in formalin and embedded in paraffin for histological analysis, while the other half was immediately frozen in liquid nitrogen and later subject to RNA extraction and expression analysis.

13.16. Statistics and plots

In all box plots, the vertical lines (whiskers) represent the maximum and minimum values, the box signifies the upper (75th) and lower quartiles (25th), the median is represented by a horizontal line within the box and the mean is represented by a rhombus or cross within the box. The statistical test and comparison used to calculate p-values is reported in each figure legend.

13.17. Antibodies table

Antibody	Source	Cat. number /Ref.	Application
macroH2A1	home-made	Buschbeck, 2009	Western Blot, ChIP, PLA
macroH2A2	home-made	Buschbeck, 2009	Western Blot, ChIP, PLA
macroH2A1.1	home-made	Sporn, 2009	Western Blot
macroH2A1.2	Cell Signalling	4827S	Western Blot
macroH2A1.2	home-made	Sporn, 2009	Western Blot (Figure 32)
H2A	Abcam	ab15653	Western Blot
H3	Abcam	ab1791	ChIP
H3K4me3	Millipore	07-473	ChIP
H3K27Ac	Abcam	ab4729	ChIP
H3K9me3	Abcam	ab8898	IF, ChIP
H3K9me3	Sigma	SAB4800018	PLA
IgG	Abcam	ab46540	ChIP
DKK1	Santa Cruz Biotechnology	sc-374574	Western Blot
FLAG	Sigma	F1804-1MG	Western Blot, IF
NPM1	Abcam	ab10530	IF
mCherry	Abcm	ab167453	Western Blot
GFP	Santa Cruz Biotechnology	sc-9996	Western Blot, IF
poly(ADP-ribose)	Trevigen	4336-APC-050	Western Blot
Lamin B1	Santa Cruz Biotechnology	sc-20682	Western Blot, PLA, IF , ChIP
AFP	R&D systems	MAB1368	IHC
β -cat	BD Biosciences	610154	IHC
Ki67	Dako	M7240	IHC (Figure 38)
Ki67	Dako	GA626	IHC (Figure 27)

13.18. Primers table

Application	Name	Sequence	References	Species
qRT-PCR	DKK1 Fwd	CCTTGAACTCGGTTCTCAATTCC		Hs
qRT-PCR	DKK1 Rev	CAATGGTCTGGTACTTATTCCCG		Hs
qRT-PCR	H2AFY Fwd	CCTGGCTGATGATAAGAAGCTG		Hs
qRT-PCR	H2AFY Fwd	GACACGAAGTAAC TGGAGATGG		Hs
qRT-PCR	H2AFY2 Fwd	CATGGCGGCAGTCATTGAG		Hs
qRT-PCR	H2AFY2 Rev	ATTGCCGGCCAATTC TAGAA		Hs
ChIP / FAIRE	CNTNAP2 Fwd	GGATGGGAGAAACAGTGGGA	(Simon, 2012)	Hs
ChIP / FAIRE	CNTNAP2 Rev	TAGGCAAGAAGGTGTGGGAG		Hs
qRT-PCR / FISH / ChIP	rDNA 45S Fwd	CGACGACCCATTCGAACGTCT	(Cong, 2013)	Hs
qRT-PCR / FISH / ChIP	rDNA 45S Rev	CTCTCCGGAATCGAACCCTGA		Hs
ChIP	rDNA 45S Fwd	GCGCAGCGTTTGCTCTCT		Mm
ChIP	rDNA 45S Rev	CACACAAGCCGAGCCACAT		Mm
qRT-PCR / ChIP / FAIRE	rDNA 5s Fwd	TCTACGGCCATACCACCCTGA	(Shen, 2013)	Hs
qRT-PCR / ChIP / FAIRE	rDNA 5s Rev	GCCTACAGCACCCGGTATTCC		Hs
qRT-PCR / ChIP / FAIRE	DXZ4 Fwd	GCCTACGTCACGCAGGAAG		Hs
qRT-PCR / ChIP / FAIRE	DXZ4 Fwd	TATGTTTGGGCAGGAAGATCG		Hs
qRT-PCR / ChIP / FAIRE	α SAT_chr4 Fwd	TCATTCCCACAACTGCGTTG	(Zeng, 2009)	Hs
qRT-PCR / ChIP / FAIRE	α SAT_chr4 Rev	TCCAACGAAGGCCACAAGA		Hs
qRT-PCR / ChIP / FISH / FAIRE	SAT2_chr1 Fwd	CATCGAATGGAAATGAAAGGAGTC	(Zeng, 2009)	Hs
qRT-PCR / ChIP / FISH / FAIRE	SAT2_chr1 Rev	ACCATTGGATGATTGCAGTCAA		Hs
qRT-PCR	GAPDH Fwd	GAGTCAACGGATTGGTTCGT		Hs
qRT-PCR	GAPDH Rev	TTGATTTTGGAGGGATCTCG		Hs
qRT-PCR	RPLP0 Fwd	TTCATTGTGGGAGCAGAC		Hs/ Mm
qRT-PCR	RPLP0 Rev	CAGCAGTTTCTCCAGAGC		Hs / Mm
ChIP / FAIRE	MBOAT7 Fwd	CTTCGGAGGTAGTCGAGTCC	(Simon , 2012)	Hs
ChIP / FAIRE	MBOAT7 Rev	CCTAGCGTCACTTGTACCCA		Hs
ChIP	Major Sat. Fwd	TGGAATATGGCGAGAAACTG	(Millanes-Romero, 2013)	Mm
ChIP	Major Sat. Rev	AGGTCCTTCAGTGGGCATTT		Mm
ChIP	Minor Sat. Fwd	GAAAATGATAAAAACCACAC	(Millanes-Romero, 2013)	Mm
ChIP	Minor Sat. Rev	ACTCATTGATATACACTGTT		Mm
ChIP	Line-L1 Fwd	TTTGGGACACAATGAAAGCA	(Martens, 2005)	Mm
ChIP	Line-L1 Rev	CTGCCGTCTACTCCTCTTGG		Mm
ChIP	E2F1 Fwd	GCGTTAAAGCCAATAGGAACC		Hs
ChIP	E2F1 Rev	AAAGTCCC GGCCACTTTTAC		Hs

Materials and Methods

ChIP	WNT8 Fwd	TGTGTGCATGTGCAATGTATG		Mm
ChIP	WNT8 Rev	CACCACTCATACCCACAGAC		Mm
ChIP	DKK1 P1 Fwd	AAAGCCGAGCAAACGGACTA		Hs
ChIP	DKK1 P1 Rev	GGACCCCCGGTCTTAATCAG		Hs
ChIP	DKK1 P2 Fwd	TCTTTGTGCGCATGGTAGCG		Hs
ChIP	DKK1 P2 Rev	GTGGGGGCAGGTTCTTGATA		Hs
ChIP	DKK1 N Fwd	TGTGTGTACGTATGTGGTCCT		Hs
ChIP	DKK1 N Rev	CCACTGGATGTGGCAGACTA		Hs
ChIP	DKK1 E1 Fwd	GGCTGGCCCTAAAAAGGACT		Hs
ChIP	DKK1 E1 Rev	GGACCAAACGTGAGGGAAGA		Hs
ChIP	DKK1 E2 Fwd	GGTTTAAGAATCAAAGATGGGGTCA		Hs
ChIP	DKK1 E2 Rev	CGACAACGTTCTAAAACAATCCCA		Hs
ChIP	DKK1 E3 Fwd	TGCCATGAGTGGCTTACAAGT		Hs
ChIP	DKK1 E3 Rev	ACTGCTTTGGTGGACACAGAT		Hs
UMI-4C	DKK1_PRM_UP	CATTTTCTCGCAGTGGCGG		Hs
UMI-4C	DKK_PRM_DN	GAGGAGGGCAACTGAAGGA		Hs

Bibliography

14. Bibliography

- Abbott, D.W., Chadwick, B.P., Thambirajah, A.A., and Ausió, J. (2005). Beyond the Xi. MacroH2A chromatin distribution and post-translational modification in an avian system. *J. Biol. Chem.* *280*, 16437–16445.
- Alipour, E., and Marko, J.F. (2012). Self-organization of domain structures by DNA-loop-extruding enzymes. *Nucleic Acids Res.* *40*, 11202–11212.
- Allen, B.L., and Taatjes, D.J. (2015). The Mediator complex: A central integrator of transcription. *Nat. Rev. Mol. Cell Biol.* *16*, 155–166.
- Allfrey, G., Faulkner, R., and Mirsky, A.E. (1964). Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *51*, 786–794.
- Allshire, R.C., and Madhani, H.D. (2017). Ten principles of heterochromatin formation and function. *Nat. Publ. Gr.* *19*, 229–244.
- Amendola, M., and van Steensel, B. (2014). Mechanisms and dynamics of nuclear lamina-genome interactions. *Curr. Opin. Cell Biol.* *28*, 61–68.
- Amendola, M., and Steensel, B. Van (2015). Nuclear lamins are not required for lamina- associated domain organization in mouse embryonic stem cells.
- Andrey, G., Montavon, T., Mascrez, B., Gonzalez, F., Noordermeer, D., Leleu, M., Trono, D., Spitz, F., and Duboule, D. (2013). A Switch Between Topological Domains Underlies HoxD Genes Collinearity in Mouse Limbs. *Science (80-.).* *340*, 1234167.
- Angelov, D., Molla, A., Perche, P., Khochbin, S., Hans, F., Co, J., Bouvet, P., Dimitrov, S., Merci, D. De, Cedex, L.T., et al. (2003). The Histone Variant MacroH2A Interferes with Transcription Factor Binding and SWI / SNF Nucleosome Remodeling. *11*, 1033–1041.
- van Attikum, H., and Gasser, S.M. (2009). Crosstalk between histone modifications during the DNA damage response. *Trends Cell Biol.* *19*, 207–217.
- Bafico, A., Liu, G., Yaniv, A., Gazit, A., and Aaronson, S.A. (2001). Novel mechanism of Wnt signalling inhibition mediated by Dickkopf-1 interaction with LRP6/Arrow. *Nat. Cell Biol.* *3*, 683–686.
- Bandarra, D., and Rocha, S. (2013). Tale of two transcription factors: NF-κB and HIF crosstalk.OA *Molecular and Cell Biology. OA Mol. Cell Biol.* *1*, 1–7.
- Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* *410*, 120–124.
- Barateau, A., and Buendia, B. (2010). In Situ Detection of Interactions Between Nuclear Envelope Proteins and Partners. *Cold Spring Harb. Perspect. Biol.* *2*, a000539–a000539.
- Barlow, J.H., Faryabi, R.B., Callén, E., Wong, N., Malhowski, A., Chen, H.T., Gutierrez-Cruz, G., Sun, H.W., McKinnon, P., Wright, G., et al. (2013). Identification of early replicating fragile sites that contribute

Bibliography

to genome instability. *Cell* 152, 620–632.

Barrero, M.J., Sese, B., Martí, M., and Belmonte, J.C.I. (2013a). Macro histone variants are critical for the differentiation of human pluripotent cells. *J. Biol. Chem.* 288, 16110–16116.

Barrero, M.J., Sese, B., Kuebler, B., Bilic, J., Boue, S., Martí, M., and Izpisua Belmonte, J.C. (2013b). Macrohistone Variants Preserve Cell Identity by Preventing the Gain of H3K4me2 during Reprogramming to Pluripotency. *Cell Rep.* 3, 1005–1011.

Barrington, C., Finn, R., and Hadjur, S. (2017). Cohesin biology meets the loop extrusion model. *Chromosom. Res.* 25, 51–60.

De Barrios, O., Gyorffy, B., Fernández-Aceñero, M.J., Sánchez-Tilló, E., Sánchez-Moral, L., Siles, L., Esteve-Arenys, A., Roué, G., Casal, J.I., Darling, D.S., et al. (2017). ZEB1-induced tumourigenesis requires senescence inhibition via activation of DKK1/mutant p53/Mdm2/CtBP and repression of macroH2A1. *Gut* 66, 666–682.

Barzily-Rokni, M., Friedman, N., Ron-Bigger, S., Isaac, S., Michlin, D., and Eden, A. (2011). Synergism between DNA methylation and macroH2A1 occupancy in epigenetic silencing of the tumor suppressor gene p16(CDKN2A). *Nucleic Acids Res.* 39, 1326–1335.

Beaudouin, J., Mora-Bermúdez, F., Klee, T., Daigle, N., and Ellenberg, J. (2006). Dissecting the contribution of diffusion and interactions to the mobility of nuclear proteins. *Biophys. J.* 90, 1878–1894.

Becker, P.B., and Hörz, W. (2002). ATP-Dependent Nucleosome Remodeling. *Annu. Rev. Biochem.* 71, 247–273.

Becker, J.S., Nicetto, D., and Zaret, K.S. (2016). H3K9me3-Dependent Heterochromatin: Barrier to Cell Fate Changes. *Trends Genet.* 32, 29–41.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell* 125, 315–326.

Bernstein, E., Muratore-Schroeder, T.L., Diaz, R.L., Chow, J.C., Changolkar, L.N., Shabanowitz, J., Heard, E., Pehrson, J.R., Hunt, D.F., and Allis, C.D. (2008). A phosphorylated subpopulation of the histone variant macroH2A1 is excluded from the inactive X chromosome and enriched during mitosis. *Proc. Natl. Acad. Sci. U. S. A.* 105, 1533–1538.

Bertout, J.A., Patel, S.A., and Simon, M.C. (2008). Hypoxia and Metabolism Series - Timeline The impact of O(2) availability on human cancer. *Nat. Rev. Cancer* 8, 967–975.

Bhattacharjee, R.N., Banks, G.C., Trotter, K.W., Lee, H.-L., and Archer, T.K. (2001). Histone H1 Phosphorylation by Cdk2 Selectively Modulates Mouse Mammary Tumor Virus Transcription through Chromatin Remodeling. *Mol. Cell. Biol.* 21, 5417–5425.

Bian, Q., Khanna, N., Alvikas, J., and Belmont, A.S. (2013). β -Globin cis-elements determine differential nuclear targeting through epigenetic modifications. *J. Cell Biol.* 203, 767–783.

Biscotti, M.A., Olmo, E., and Heslop-Harrison, J.S. (Pat. (2015). Repetitive DNA in eukaryotic genomes.

Chromosom. Res. 23, 415–420.

Biterge, B., and Schneider, R. (2014). Histone variants: Key players of chromatin. *Cell Tissue Res.* 356, 457–466.

Black, B.E., Jansen, L.E.T., Maddox, P.S., Foltz, D.R., Desai, A.B., Shah, J. V., and Cleveland, D.W. (2007). Centromere Identity Maintained by Nucleosomes Assembled with Histone H3 Containing the CENP-A Targeting Domain. *Mol. Cell* 25, 309–322.

Blackledge, N.P., Farcas, A.M., Kondo, T., King, H.W., McGouran, J.F., Hanssen, L.L.P., Ito, S., Cooper, S., Kondo, K., Koseki, Y., et al. (2014). Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. *Cell* 157, 1445–1459.

Blackledge, N.P., Rose, N.R., and Klose, R.J. (2015). Targeting Polycomb systems to regulate gene expression: Modifications to a complex story. *Nat. Rev. Mol. Cell Biol.* 16, 643–649.

Bolland, D.J., King, M.R., Reik, W., Corcoran, A.E., and Krueger, C. (2013). Robust 3D DNA FISH using directly labeled probes. *J. Vis. Exp.* 1–9.

Bonev, B., Mendelson Cohen, N., Szabo, Q., Fritsch, L., Papadopoulos, G.L., Lubling, Y., Xu, X., Lv, X., Hugnot, J.P., Tanay, A., et al. (2017). Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell* 171, 557-572.e24.

Bönisch, C., and Hake, S.B. (2012). Histone H2A variants in nucleosomes and chromatin: More or less stable? *Nucleic Acids Res.* 40, 10719–10741.

Bönisch, C., Schneider, K., Pünzeler, S., Wiedemann, S.M., Bielmeier, C., Bocola, M., Eberl, H.C., Kuegel, W., Neumann, J., Kremmer, E., et al. (2012). H2A.Z.2.2 is an alternatively spliced histone H2A.Z variant that causes severe nucleosome destabilization. *Nucleic Acids Res.* 40, 5951–5964.

Bowerman, S., Hickok, R.J., and Wereszczynski, J. (2019). Unique Dynamics in Asymmetric macroH2A-H2A Hybrid Nucleosomes Result in Increased Complex Stability. *J. Phys. Chem. B* 123, 419–427.

Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K., et al. (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* 441, 349–353.

Boyle, S., Gilchrist, S., Bridger, J.M., Mahy, N.L., Ellis, J.A., and Bickmore, W.A. (2001). The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Hum. Mol. Genet.* 10, 211–219.

Branco, M.R., and Pombo, A. (2006). Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. *PLoS Biol.* 4, 780–788.

Bronshtein, I., Kanter, I., Kepten, E., Lindner, M., Berezin, S., Shav-Tal, Y., and Garini, Y. (2016). Exploring chromatin organization mechanisms through its dynamic properties. *Nucleus* 7, 27–33.

Bulut-Karslioglu, A., Perrera, V., Scaranaro, M., De La Rosa-Velazquez, I.A., Van De Nobelen, S., Shukeir, N., Popow, J., Gerle, B., Opravil, S., Pagani, M., et al. (2012). A transcription factor-based mechanism for mouse heterochromatin formation. *Nat. Struct. Mol. Biol.* 19, 1023–1032.

Bibliography

- Buschbeck, M., and Hake, S.B. (2017). Variants of core histones and their roles in cell fate decisions, development and cancer. *Nat. Rev. Mol. Cell Biol.* **18**, 299–314.
- Buschbeck, M., Uribesalzo, I., Wibowo, I., Rué, P., Martin, D., Gutierrez, A., Morey, L., Guigó, R., López-Schier, H., and Di Croce, L. (2009). The histone variant macroH2A is an epigenetic regulator of key developmental genes. *Nat. Struct. Mol. Biol.* **16**, 1074–1079.
- Cairo, S., Armengol, C., De Reyniès, A., Wei, Y., Thomas, E., Renard, C.A., Goga, A., Balakrishnan, A., Semeraro, M., Gresh, L., et al. (2008). Hepatic Stem-like Phenotype and Interplay of Wnt/ β -Catenin and Myc Signaling in Aggressive Childhood Liver Cancer. *Cancer Cell* **14**, 471–484.
- Caneparo, L., Huang, Y.-L., Staudt, N., Tada, M., Ahrendt, R., Kazanskaya, O., Niehrs, C., and Houart, C. (2007). Dickkopf-1 regulates gastrulation movements by coordinated modulation of Wnt/ β -catenin and Wnt/PCP activities, through interaction with the Dally-like homolog Knypek. *Genes Dev.* **21**, 465–480.
- Cantariño, N., Douet, J., and Buschbeck, M. (2013). MacroH2A - An epigenetic regulator of cancer. *Cancer Lett.* **336**, 247–252.
- Cantariño, N., Teresa Fernández-Figueras, M., Valero, V., Musulén, E., Malinverni, R., Granada, I., Goldie, S.J., Martín-Caballero, J., Douet, J., Forcales, S.V., et al. (2016). A cellular model reflecting the phenotypic heterogeneity of mutant HRAS driven squamous cell carcinoma. *Int. J. Cancer* **139**, 1106–1116.
- Canzio, D., Chang, E.Y., Shankar, S., Kuchenbecker, K.M., Simon, M.D., Madhani, H.D., Narlikar, G.J., and Al-Sady, B. (2011). Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. *Mol. Cell* **41**, 67–81.
- Cao, R., and Zhang, Y. (2004). SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. *Mol. Cell* **15**, 57–67.
- Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in polycomb-group silencing. *Science* (80-.). **298**, 1039–1043.
- Cao, R., Tsukada, Y., Zhang, Y., Hill, C., and Carolina, N. (2005). Role of Bmi-1 and Ring1A in H2A Ubiquitylation and Hox Gene Silencing. *Cell* **120**, 845–854.
- Casper, J., Zweig, A.S., Villarreal, C., Tyner, C., Speir, M.L., Rosenbloom, K.R., Raney, B.J., Lee, C.M., Lee, B.T., Karolchik, D., et al. (2018). The UCSC Genome Browser database: 2018 update. *Nucleic Acids Res.* **46**, D762–D769.
- Caterino, T.L., Fang, H., and Hayes, J.J. (2011). Nucleosome Linker DNA Contacts and Induces Specific Folding of the Intrinsically Disordered H1 Carboxyl-Terminal Domain. *Mol. Cell. Biol.* **31**, 2341–2348.
- Cesarini, E., Mozzetta, C., Marullo, F., Gregoret, F., Gargiulo, A., Columbaro, M., Cortesi, A., Antonelli, L., Di Pelino, S., Squarzon, S., et al. (2015). Lamin A/C sustains PcG protein architecture, maintaining transcriptional repression at target genes. *J. Cell Biol.* **211**, 533–551.
- Chadwick, B.P. (2008). DXZ4 chromatin adopts an opposing conformation to that of the surrounding chromosome and acquires a novel inactive X-specific role involving CTCF and antisense transcripts. *Chromosoma* **117**, 162–172.

1259–1269.

Chadwick, B.P., and Willard, H.F. (2001). Histone H2A variants and the inactive X chromosome: identification of a second macroH2A variant. *Hum. Mol. Genet.* 10, 1101–1113.

Chadwick, B.P., Valley, C.M., and Willard, H.F. (2001). Histone variant macroH2A contains two distinct macrochromatin domains capable of directing macroH2A to the inactive X chromosome. *Nucleic Acids Res.* 29, 2699–2705.

Chakravarthy, S., and Luger, K. (2006). The histone variant macro-H2A preferentially forms “hybrid nucleosomes.” *J. Biol. Chem.* 281, 25522–25531.

Chakravarthy, S., Gundimella, S.K.Y., Caron, C., Perche, P., Pehrson, J.R., and Khochbin, S. (2005a). Structural Characterization of the Histone Variant macroH2A. *Mol. Cell. Biol.* 25, 7616–7624.

Chakravarthy, S., Gundimella, S.K.Y., Caron, C., Perche, P.-Y., Pehrson, J.R., Khochbin, S., and Luger, K. (2005b). Structural Characterization of the Histone Variant macroH2A. *Mol. Cell. Biol.* 25, 7616–7624.

Chakravarthy, S., Patel, A., and Bowman, G.D. (2012). The basic linker of macroH2A stabilizes DNA at the entry/exit site of the nucleosome. *Nucleic Acids Res.* 40, 8285–8295.

Chang, E.Y., Ferreira, H., Somers, J., Nusinow, A., Owen-hughes, T., Narlikar, G.J., and Nusinow, D.A. (2008). MacroH2A Allows ATP-Dependent Chromatin Remodeling by SWI / SNF and ACF Complexes but Specifically Reduces Recruitment of SWI / SNF. *Biochemistry* 47, 13726–13732.

Changolkar, L.N., and Pehrson, J.R. (2002). Reconstitution of nucleosomes with histone macroH2A1.2. *Biochemistry* 41, 179–184.

Changolkar, L.N., and Pehrson, J.R. (2006). macroH2A1 Histone Variants Are Depleted on Active Genes but Concentrated on the Inactive X Chromosome. *Mol. Cell. Biol.* 26, 4410–4420.

Changolkar, L.N., Costanzi, C., Leu, N.A., Chen, D., McLaughlin, K.J., and Pehrson, J.R. (2007). Developmental changes in histone macroH2A1-mediated gene regulation. *Mol. Cell. Biol.* 27, 2758–2764.

Changolkar, L.N., Singh, G., Cui, K., Berletch, J.B., Zhao, K., Disteché, C.M., and Pehrson, J.R. (2010). Genome-Wide Distribution of MacroH2A1 Histone Variants in Mouse Liver Chromatin. *Mol. Cell. Biol.* 30, 5473–5483.

Chen, H., Ruiz, P.D., Novikov, L., Casill, A.D., Park, J.W., and Gamble, M.J. (2014). MacroH2A1.1 and PARP-1 cooperate to regulate transcription by promoting CBP-mediated H2B acetylation. *Nat. Struct. Mol. Biol.* 21, 981–989.

Chen, H., Ruiz, P.D., McKimpson, W.M., Novikov, L., Kitsis, R.N., and Gamble, M.J. (2015). MacroH2A1 and ATM Play Opposing Roles in Paracrine Senescence and the Senescence-Associated Secretory Phenotype. *Mol. Cell* 59, 719–731.

Chen, L., Li, M., Li, Q., Wang, C., and Xie, S. (2013). DKK1 promotes hepatocellular carcinoma cell migration and invasion through β -catenin/MMP7 signaling pathway. *Mol. Cancer* 12, 157.

Chen, W., Zhang, Y.W., Li, Y., Zhang, J.W., Zhang, T., Fu, B.S., Zhang, Q., and Jiang, N. (2016). Constitutive expression of Wnt/ β -catenin target genes promotes proliferation and invasion of liver cancer

Bibliography

stem cells. *Mol. Med. Rep.* 13, 3466–3474.

Cong, R., Das, S., Douet, J., Wong, J., Buschbeck, M., Mongelard, F., and Bouvet, P. (2014). MacroH2A1 histone variant represses rDNA transcription. *Nucleic Acids Res.* 42, 181–192.

Contreras, A., Hale, T.K., Stenoien, D.L., Rosen, J.M., Mancini, M.A., and Herrera, R.E. (2003). The Dynamic Mobility of Histone H1 Is Regulated by Cyclin/CDK Phosphorylation. *Mol. Cell. Biol.* 23, 8626–8636.

Corujo, D., and Buschbeck, M. (2018). Post-translational modifications of H2A histone variants and their role in cancer. *Cancers (Basel)*. 1–25.

Costanzi, C., and Pehrson, J.R. (1998). Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature* 393, 599–601.

Costanzi, C., and Pehrson, J.R. (2001). MACROH2A2, a New Member of the MACROH2A Core Histone Family. *J. Biol. Chem.* 276, 21776–21784.

Cowieson, N.P., Partridge, J.F., Allshire, R.C., and McLaughlin, P.J. (2000). Dimerisation of a chromo shadow domain and distinctions from the chromodomain as revealed by structural analysis. *Curr. Biol.* 10, 517–525.

Cremer, T., and Cremer, M. (2010). Chromosome territories. *Cold Spring Harb. Perspect. Biol.* 2, 1–23.

Cremer, M., Hase, J. V., Volm, T., Brero, A., Kreth, G., Walter, J., Fischer, C., Solovei, I., Cremer, C., and Cremer, T. (2001). Non-random radial higher-order chromatin arrangements in nuclei of diploid human cells. *Chromosom. Res.* 9, 541–567.

Creppe, C., Janich, P., Cantarino, N., Noguera, M., Valero, V., Musulen, E., Douet, J., Posavec, M., Martin-Caballero, J., Sumoy, L., et al. (2012). MacroH2A1 Regulates the Balance between Self-Renewal and Differentiation Commitment in Embryonic and Adult Stem Cells. *Mol. Cell. Biol.* 32, 1442–1452.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U. S. A.* 107, 21931–21936.

Di Croce, L., and Helin, K. (2013). Transcriptional regulation by Polycomb group proteins. *Nat. Struct. Mol. Biol.* 20, 1147–1155.

Croft, J.A., Bridger, J.M., Boyle, S., Perry, P., Teague, P., and Bickmore, W.A. (1999). Differences in the localization and morphology of chromosomes in the human nucleus. *J. Cell Biol.* 145, 1119–1131.

Cummins, E.P., and Taylor, C.T. (2017). Hypoxia and inflammation. *Biochem. (Lond)*. 39, 34–36.

Danan-Gotthold, M., Golan-Gerstl, R., Eisenberg, E., Meir, K., Karni, R., and Levanon, E.Y. (2015). Identification of recurrent regulated alternative splicing events across human solid tumors. *Nucleic Acids Res.* 43, 5130–5144.

Dardenne, E., Pierredon, S., Driouch, K., Gratadou, L., Lacroix-triki, M., Espinoza, M.P., Zonta, E., Germann, S., Mortada, H., Villemin, J., et al. (2012). Splicing switch of an epigenetic regulator by RNA helicases promotes tumor-cell invasiveness. *Nat. Struct. Mol. Biol.* 19, 1139–1146.

- Dechat, T., Gesson, K., and Foisner, R. (2010). Lamina-independent lamins in the nuclear interior serve important functions. *Cold Spring Harb. Symp. Quant. Biol.* 75, 533–543.
- Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. *Science* (80-.). 295, 1306–1311.
- Deng, W., Lee, J., Wang, H., Miller, J., Reik, A., Gregory, P.D., Dean, A., and Blobel, G.A. (2012). Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* 149, 1233–1244.
- Deribe, Y.L., Pawson, T., and Dikic, I. (2010). Post-translational modifications in signal integration. *Nat. Struct. Mol. Biol.* 17, 666–672.
- Désert, R., Mebarki, S., Desille, M., Sicard, M., Lavergne, E., Renaud, S., Bergeat, D., Sulpice, L., Perret, C., Turlin, B., et al. (2016). “Fibrous nests” in human hepatocellular carcinoma express a Wnt-induced gene signature associated with poor clinical outcome. *Int. J. Biochem. Cell Biol.* 81, 195–207.
- Dialynas, G.K., Vitalini, M.W., and Wallrath, L.L. (2008). Linking Heterochromatin Protein 1 (HP1) to cancer progression. *Mutat. Res. - Fundam. Mol. Mech. Mutagen.* 647, 13–20.
- Diarra, D., Stolina, M., Polzer, K., Zwerina, J., Ominsky, M.S., Dwyer, D., Korb, A., Smolen, J., Hoffmann, M., Scheinecker, C., et al. (2007). Dickkopf-1 is a master regulator of joint remodeling. *13*.
- Diesch, J., Hannan, R.D., and Sanij, E. (2014). Perturbations at the ribosomal genes loci are at the centre of cellular dysfunction and human disease. 32–41.
- Dillinger, S., Straub, T., and Németh, A. (2017). Nucleolus association of chromosomal domains is largely maintained in cellular senescence despite massive nuclear reorganisation. *PLoS One* 12, e0178821.
- Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376–380.
- Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W., et al. (2015). Chromatin architecture reorganization during stem cell differentiation. *Nature* 518, 331–336.
- Dixon, J.R., Gorkin, D.U., and Ren, B. (2016). Chromatin Domains: The Unit of Chromosome Organization. *Mol. Cell* 62, 668–680.
- Dobrzynska, A., Gonzalo, S., Shanahan, C., and Askjaer, P. (2016). The Nuclear Lamina in Health and Disease. *Nucleus* 7, 0.
- Douet, J., Corujo, D., Malinverni, R., Renaud, J., Sansoni, V., Posavec-Marjanović, M., Cantariño, N., Valero, V., Mongelard, F., Bouvet, P., et al. (2017). MacroH2A histone variants maintain nuclear organization and heterochromatin architecture. *J. Cell Sci.* jcs.199216.
- Downen, J.M., Fan, Z.P., Hnisz, D., Ren, G., Abraham, B.J., Zhang, L.N., Weintraub, A.S., Schuijers, J., Lee, T.I., Zhao, K., et al. (2014). Control of Cell Identity Genes Occurs in Insulated Neighborhoods in Mammalian Chromosomes. *Cell* 159, 374–387.

Bibliography

- Doyen, C., An, W., Angelov, D., Bondarenko, V., Mietton, F., Studitsky, V.M., Hamiche, A., Roeder, R.G., Bouvet, P., and Dimitrov, S. (2006). Mechanism of Polymerase II Transcription Repression by the Histone Variant macroH2A. *Mol. Cell. Biol.* 26, 1156–1164.
- ENCODE Project Consortium, Dunham, I., Kundaje, A., Aldred, S.F., Collins, P.J., Davis, C. a, Doyle, F., Epstein, C.B., Fietze, S., Harrow, J., et al. (2013). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74.
- Endo, Y., Beauchamp, E., Woods, D., Taylor, W.G., Toretsky, J.A., Uren, A., and Rubin, J.S. (2008). Wnt-3a and Dickkopf-1 Stimulate Neurite Outgrowth in Ewing Tumor Cells via a Frizzled3- and c-Jun N-Terminal Kinase-Dependent Mechanism. *Mol. Cell. Biol.* 28, 2368–2379.
- Endoh, M., Endo, T.A., Endoh, T., Fujimura, Y.I., Ohara, O., Toyoda, T., Otte, A.P., Okano, M., Brockdorff, N., Vidal, M., et al. (2008). Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity. *Development* 135, 1513–1524.
- Entrevan, M., Schuettengruber, B., and Cavalli, G. (2016). Regulation of Genome Architecture and Function by Polycomb Proteins. *Trends Cell Biol.* 26, 511–525.
- Ernst, J., and Kellis, M. (2010). Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat. Biotechnol.* 28, 817–838.
- Falandry, C., Fourel, G., Galy, V., Ristriani, T., Horard, B., Bensimon, E., Salles, G., Gilson, E., and Magdinier, F. (2010). CLLD8/KMT1F is a lysine methyltransferase that is important for chromosome segregation. *J. Biol. Chem.* 285, 20234–20241.
- Fatima, S. (2011). Dickkopfs and Wnt/ β -catenin signalling in liver cancer. *World J. Clin. Oncol.* 2, 311.
- Fawcett, D.W. (1966). On the occurrence of a fibrous lamina on the inner aspect of the nuclear envelope in certain cells of vertebrates. *Am. J. Anat.* 119, 129–145.
- Feinberg, A.P., Koldobskiy, M.A., and Göndör, A. (2016). Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. *Nat. Rev. Genet.* 17, 284–299.
- Finger, E.C., and Giaccia, A.J. (2010). Hypoxia, inflammation, and the tumor microenvironment in metastatic disease. *Cancer Metastasis Rev.* 29, 285–293.
- Finlan, L.E., Sproul, D., and Thomson, I. (2008). Recruitment to the nuclear periphery can alter expression of genes in human cells. *PLoS Genet.* 21, 89–90.
- Fouad, Y.M., Mohamed, H.I., Enas, M., and Rasek, M.A. (2016). Clinical significance and diagnostic value of serum dickkopf-1 in patients with hepatocellular carcinoma. 5521.
- Francis, N.J., Kingston, R.E., and Woodcock, C.L. (2004). Chromatin compaction by a polycomb group protein complex. *Science* (80-). 306, 1574–1577.
- Franke, M., Ibrahim, D.M., Andrey, G., Schwarzer, W., Heinrich, V., Schöpflin, R., Kraft, K., Kempfer, R., Jerković, I., Chan, W.L., et al. (2016). Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature* 538, 265–269.
- Fraser, P., and Bickmore, W. (2007). Nuclear organization of the genome and the potential for gene

regulation. *Nature* 447, 413–417.

Freund, A., Laberge, R.M., Demaria, M., and Campisi, J. (2012). Lamin B1 loss is a senescence-associated biomarker. *Mol. Biol. Cell* 23, 2066–2075.

Fu, Y., Lv, P., Yan, G., Fan, H., Cheng, L., Zhang, F., Dang, Y., Wu, H., and Wen, B. (2015). MacroH2A1 associates with nuclear lamina and maintains chromatin architecture in mouse liver cells. *Sci. Rep.* 5, 17186.

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., and Mirny, L.A. (2016). Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep.* 15, 2038–2049.

Fujisawa, T., and Filippakopoulos, P. (2017). Functions of bromodomain-containing proteins and their roles in homeostasis and cancer. *Nat. Rev. Mol. Cell Biol.* 18, 246–262.

Furlong, E.E.M., and Levine, M. (2018). Developmental enhancers and chromosome topology. *Science* (80-.). 361, 1341–1345.

Fussner, E., Strauss, M., Djuric, U., Li, R., Ahmed, K., Hart, M., Ellis, J., and Bazett-Jones, D.P. (2012). Open and closed domains in the mouse genome are configured as 10-nm chromatin fibres. *EMBO Rep.* 13, 992–996.

Gamble, M.J., and Kraus, W.L. (2010). Multiple facets of the unique histone variant macroH2A: From genomics to cell biology. *Cell Cycle* 9, 2568–2574.

Gamble, M.J., Frizzell, K.M., Yang, C., Krishnakumar, R., and Kraus, W.L. (2010). The histone variant macroH2A1 marks repressed autosomal chromatin, but protects a subset of its target genes from silencing. *Genes Dev.* 24, 21–32.

García-Cao, M., O'Sullivan, R., Peters, A.H.F.M., Jenuwein, T., and Blasco, M.A. (2004). Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases. *Nat. Genet.* 36, 94–99.

Gaspar-Maia, A., Qadeer, Z. a, Hasson, D., Ratnakumar, K., Leu, N.A., Leroy, G., Liu, S., Costanzi, C., Valle-Garcia, D., Schaniel, C., et al. (2013). MacroH2A histone variants act as a barrier upon reprogramming towards pluripotency. *Nat. Commun.* 4, 1565.

Gel, B., Di, A., Serra, E., Buschbeck, M., Peinado, M.A., and Malinverni, R. (2016). Genome analysis regioneR : an R / Bioconductor package for the association analysis of genomic regions based on permutation tests. 32, 289–291.

Gilbert, D.M. (2002). Replication timing and transcriptional control: Beyond cause and effect - part III. *Curr. Opin. Cell Biol.* 14, 377–383.

Gilkes, D.M., Semenza, G.L., and Wirtz, D. (2014). Hypoxia and the extracellular matrix: Drivers of tumour metastasis. *Nat. Rev. Cancer* 14, 430–439.

Glover, T.W., Wilson, T.E., and Arlt, M.F. (2017). Fragile sites in cancer: more than meets the eye. *Nat. Rev. Cancer* 17, 489–501.

Goldberg, A.D., Allis, C.D., and Bernstein, E. (2007). Epigenetics: A Landscape Takes Shape. *Cell* 128,

Bibliography

635–638.

Gonzalez-Sandoval, A., Towbin, B.D., Kalck, V., Cabianca, D.S., Gaidatzis, D., Hauer, M.H., Geng, L., Wang, L., Yang, T., Wang, X., et al. (2015). Perinuclear Anchoring of H3K9-Methylated Chromatin Stabilizes Induced Cell Fate in *C. elegans* Embryos. *Cell* 163, 1333–1347.

Grau, D.J., Chapman, B.A., Garlick, J.D., Borowsky, M., Francis, N.J., and Kingston, R.E. (2011). Compaction of chromatin by diverse polycomb group proteins requires localized regions of high charge. *Genes Dev.* 25, 2210–2221.

Grigoryev, S.A., and Woodcock, C.L. (2012). Chromatin organization - The 30nm fiber. *Exp. Cell Res.* 318, 1448–1455.

Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W., et al. (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* 453, 948–951.

Guo, Y., Xu, Q., Canzio, D., Krainer, A.R., Maniatis, T., Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., et al. (2015). CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer / Promoter Function. *Cell* 162, 900–910.

Gurard-Levin, Z.A., Quivy, J.-P., and Almouzni, G. (2014). Histone Chaperones: Assisting Histone Traffic and Nucleosome Dynamics. *Annu. Rev. Biochem.* 83, 487–517.

Haarhuis, J.H.I., Van Der Weide, R.H., Blomen, V.A., Brummelkamp, T.R., De Wit, E., and Rowland, B.D. (2017). The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. 693–707.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: The next generation. *Cell* 144, 646–674.

Harikumar, A., and Meshorer, E. (2015a). Chromatin remodeling and bivalent histone modifications in embryonic stem cells. *EMBO Rep.* 16, 1609–1619.

Harikumar, A., and Meshorer, E. (2015b). Chromatin remodeling and bivalent histone modifications in embryonic stem cells. *EMBO Rep.* 16, 1609–1619.

Harr, J.C., Luperchio, T.R., Wong, X., Cohen, E., Wheelan, S.J., and Reddy, K.L. (2015). Directed targeting of chromatin to the nuclear lamina is mediated by chromatin state and A-type lamins. *J. Cell Biol.* 208, 33–52.

Herbert, A.D., Carr, A.M., Hoffmann, E., and Lichten, M. (2014). FindFoci: A focus detection algorithm with automated parameter training that closely matches human assignments, reduces human inconsistencies and increases speed of analysis. *PLoS One* 9.

Hergeth, S.P., and Schneider, R. (2015). The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. *EMBO Rep.* 16, 1439–1453.

Hernández-Muñoz, I., Lund, A.H., van der Stoop, P., Boutsma, E., Muijters, I., Verhoeven, E., Nusinow, D.A., Panning, B., Marahrens, Y., and van Lohuizen, M. (2005). Stable X chromosome inactivation involves the PRC1 Polycomb complex and requires histone MACROH2A1 and the CULLIN3/SPOP ubiquitin E3 ligase. *Proc. Natl. Acad. Sci. U. S. A.* 102, 7635–7640.

- Hirano, Y., Hizume, K., Kimura, H., Takeyasu, K., Haraguchi, T., and Hiraoka, Y. (2012). Lamin B receptor recognizes specific modifications of histone H4 in heterochromatin formation. *J. Biol. Chem.* 287, 42654–42663.
- Hodge, D.Q., Cui, J., Gamble, M.J., and Guo, W. (2018). Histone Variant MacroH2A1 Plays an Isoform-Specific Role in Suppressing Epithelial-Mesenchymal Transition. *Sci. Rep.* 8, 1–11.
- Hoghoughi, N., Barral, S., Vargas, A., Rousseaux, S., and Khochbin, S. (2017). Histone variants: essential actors in male genome programming. *J. Biochem.* 163, 97–103.
- Holliday, R., and Pugh, J.E. (1975). DNA Modification Mechanisms and Gene Activity during Development Developmental clocks may depend on the enzymic modification of specific bases in repeated DNA sequences. *Science* (80-.). 187, 226–232.
- Hong, L., Schroth, G.P., Matthews, H.R., Yau, P., and Bradbury, E.M. (1993). Studies of the DNA binding properties of histone H4 amino terminus. *J. Biol. Chem.* 268, 305–314.
- Hu, W.-H., Miyai, K., Sporn, J.C., Luo, L., Wang, J.Y.J., Cosman, B., and Ramamoorthy, S. (2016). Loss of histone variant macroH2A2 expression associates with progression of anal neoplasm. *J. Clin. Pathol.* 69, 627–631.
- Hu, Z., Van Rooijen, N., and Yang, Y.G. (2011). Macrophages prevent human red blood cell reconstitution in immunodeficient mice. *Blood* 118, 5938–5946.
- Huang, Y., Yang, X., Zhao, F., Shen, Q., Wang, Z., Lv, X., Hu, B., Yu, B., Fan, J., and Qin, W. (2014). Overexpression of Dickkopf-1 predicts poor prognosis for patients with hepatocellular carcinoma after orthotopic liver transplantation by promoting cancer metastasis and recurrence. *Med. Oncol.* 31, 966.
- Huynh, L.M., Shinagawa, T., and Ishii, S. (2016). Two Histone Variants TH2A and TH2B Enhance Human Induced Pluripotent Stem Cell Generation. *Stem Cells Dev.* 25, 251–258.
- Illingworth, R.S., and Bird, A.P. (2009). CpG islands--'a rough guide'. *FEBS Lett.* 583, 1713–1720.
- Illingworth, R.S., Moffat, M., Mann, A.R., Read, D., Hunter, C.J., Pradeepa, M.M., Adams, I.R., and Bickmore, W.A. (2015). The E3 ubiquitin ligase activity of RING1B is not essential for early mouse development. *Genes Dev.* 29, 1897–1902.
- Iwama, A. (2019). Blood Spotlight Polycomb repressive complexes in hematological malignancies. 130, 23–30.
- Jang, E.S., Jeong, S., Kim, J., Choi, Y.S., Leissner, P., and Brechot, C. (2016). Diagnostic Performance of Alpha- Fetoprotein , Protein Induced by Vitamin K Absence , Osteopontin , Dickkopf-1 and Its Combinations for Hepatocellular Carcinoma. *PLoS One* 11, 1–14.
- Jankevicius, G., Hassler, M., Golia, B., Rybin, V., Zacharias, M., Timinszky, G., and Ladurner, A.G. (2013). A family of macrodomain proteins reverses cellular mono-ADP-ribosylation. *Nat. Struct. Mol. Biol.* 20, 508–514.
- Jansen, A., and Verstrepen, K.J. (2011). Nucleosome Positioning in *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* 75, 301–320.

Bibliography

- Janssen, A., Colmenares, S.U., and Karpen, G.H. (2018). Heterochromatin: Guardian of the Genome. *Annu. Rev. Cell Dev. Biol.* 34, 265–288.
- Javierre, B.M., Burren, O.S., Wilder, S.P., Kreuzhuber, R., Hill, S.M., Sewitz, S., Cairns, J., Wingett, S.W., Várnai, C., Thiecke, M.J., et al. (2016). Lineage-Specific Genome Architecture Links Enhancers and Non-coding Disease Variants to Target Gene Promoters. *Cell* 167, 1369-1384.e19.
- Jiao, L., and Liu, X. (2015). Structural basis of histone H3K27 trimethylation by an active polycomb repressive complex 2. *Science* 350, aac4383.
- Jin, F., Li, Y., Dixon, J.R., Selvaraj, S., Ye, Z., Lee, A.Y., Yen, C.-A., Schmitt, A.D., Espinoza, C. a, and Ren, B. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* 503, 290–294.
- Jones, P.A. (2012). Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* 13, 484–492.
- Jones, P.A., and Baylin, S.B. (2007). The epigenomics of cancer. *Cell* 683–692.
- Jones, P.A., Issa, J.P.J., and Baylin, S. (2016). Targeting the cancer epigenome for therapy. *Nat. Rev. Genet.* 17, 630–641.
- Jorgensen, S., Schotta, G., and Sorensen, C.S. (2013). Histone H4 Lysine 20 methylation: Key player in epigenetic regulation of genomic integrity. *Nucleic Acids Res.* 41, 2797–2806.
- Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* 26, 795–800.
- Kagey, M.H., and He, X. (2017). Rationale for targeting the Wnt signalling modulator Dickkopf-1 for oncology. *Br. J. Pharmacol.* 174, 4637–4650.
- Kapoor, A., Goldberg, M.S., Cumberland, L.K., Ratnakumar, K., Segura, M.F., Emanuel, P.O., Menendez, S., Vardabasso, C., Leroy, G., Vidal, C.I., et al. (2010). The histone variant macroH2A suppresses melanoma progression through regulation of CDK8. *Nature* 468, 1105–1109.
- El Kennani, S., Adrait, A., Shaytan, A.K., Khochbin, S., Bruley, C., Panchenko, A.R., Landsman, D., Pflieger, D., and Govin, J. (2017). MS-HistoneDB, a manually curated resource for proteomic analysis of human and mouse histones. *Epigenetics and Chromatin* 10, 1–18.
- Khurana, S., Kruhlak, M.J., Kim, J., Tran, A.D., Liu, J., Nyswaner, K., Shi, L., Jailwala, P., Sung, M.H., Hakim, O., et al. (2014). A macrohistone variant links dynamic chromatin compaction to BRCA1-dependent genome maintenance. *Cell Rep.* 8, 1049–1062.
- Kim, J.-M., Heo, K., Choi, J., Kim, K., and An, W. (2013). The histone variant MacroH2A regulates Ca²⁺ influx through TRPC3 and TRPC6 channels. *Oncogenesis* 2, 1–9.
- Kim, J., Sturgill, D., Sebastian, R., Khurana, S., Tran, A.D., Edwards, G.B., Kruswick, A., Burkett, S., Hosogane, E.K., Hannon, W.W., et al. (2018a). Replication Stress Shapes a Protective Chromatin Environment across Fragile Genomic Regions. *Mol. Cell* 69, 36-47.e7.
- Kim, J., Shin, Y., Lee, S., Kim, M., Punj, V., Lu, J.F., Shin, H., Kim, K., Ulmer, T.S., Koh, J., et al. (2018b).

Regulation of Breast Cancer-Induced Osteoclastogenesis by MacroH2A1.2 Involving EZH2-Mediated H3K27me3. *Cell Rep.* 24, 224–237.

Kim, S.U., Park, J.H., Kim, H.S., Lee, J.M., Lee, H.G., Kim, H., Choi, S.H., Baek, S., Kim, B.K., Park, J.Y., et al. (2015). Serum dickkopf-1 as a biomarker for the diagnosis of hepatocellular carcinoma. *Yonsei Med. J.* 56, 1296–1306.

Kind, J., Pagie, L., Ortabozkoyun, H., Boyle, S., de Vries, S.S., Janssen, H., Amendola, M., Nolen, L.D., Bickmore, W.A., and van Steensel, B. (2013a). Single-Cell Dynamics of Genome-Nuclear Lamina Interactions. *Cell* 153, 178–192.

Kind, J., Pagie, L., Ortabozkoyun, H., Boyle, S., de Vries, S.S., Janssen, H., Amendola, M., Nolen, L.D., Bickmore, W.A., and van Steensel, B. (2013b). Single-Cell Dynamics of Genome-Nuclear Lamina Interactions. *Cell* 153, 178–192.

Kind, J., Pagie, L., De Vries, S.S., Nahidiazar, L., Dey, S.S., Bienko, M., Zhan, Y., Lajoie, B., De Graaf, C.A., Amendola, M., et al. (2015). Genome-wide Maps of Nuclear Lamina Interactions in Single Human Cells. *Cell* 163, 134–147.

Koningsbruggen, S. Van, Gierlin, M., Martin, D., Barton, G.J., Ariyurek, Y., Dunnen, J.T. Den, and Lamond, A.I. (2010). High-Resolution Whole-Genome Sequencing Reveals That Specific Chromatin Domains from Most Human Chromosomes Associate with Nucleoli. *21*, 3735–3748.

Kozłowski, M., Corujo, D., Hothorn, M., Guberovic, I., Mandemaker, I.K., Blessing, C., Sporn, J., Gutierrez-Triana, A., Smith, R., Portmann, T., et al. (2018). MacroH2A histone variants limit chromatin plasticity through two distinct mechanisms. *EMBO Rep.* 19, 1–13.

Krepler, C., Wacheck, V., Strommer, S., Hartmann, G., Polterauer, P., Wolff, K., Pehamberger, H., and Jansen, B. (2004). CpG Oligonucleotides Elicit Antitumor Responses in a Human Melanoma NOD/SCID Xenotransplantation Model. *J. Invest. Dermatol.* 122, 387–391.

Krijger, P.H.L., and de Laat, W. (2016). Regulation of disease-associated gene expression in the 3D genome. *Nat. Rev. Mol. Cell Biol.* 17, 771–782.

Kruhlak, M.J., Celeste, A., Dellaire, G., Fernandez-capetillo, O., Müller, W.G., McNally, J.G., Bazett-jones, D.P., and Nussenzweig, A. (2006). Changes in chromatin structure and mobility in living cells at sites of DNA double-strand breaks. *172*, 823–834.

Kumaran, R.I., and Spector, D.L. (2008). A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. *J. Cell Biol.* 180, 51–65.

Kustatscher, G., Hothorn, M., Scheffzek, K., and Ladurner, A.G. (2005). Splicing regulates NAD metabolite binding to histone macroH2A. *Nat. Struct. Mol. Biol.* 12, 624–625.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410, 116–120.

Lacoste, N., Woolfe, A., Tachiwana, H., Garea, A.V., Barth, T., Cantaloube, S., Kurumizaka, H., Imhof, A., and Almouzni, G. (2014). Mislocalization of the Centromeric Histone Variant CenH3/CENP-A in Human Cells Depends on the Chaperone DAXX. *Mol. Cell* 53, 631–644.

Bibliography

- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359.
- Larson, A.G., and Narlikar, G.J. (2018). The Role of Phase Separation in Heterochromatin Formation, Function, and Regulation. *Biochemistry* 57, 2540–2548.
- Larson, A.G., Elnatan, D., Keenen, M.M., Trnka, M.J., Johnston, J.B., Burlingame, A.L., Agard, D.A., Redding, S., and Narlikar, G.J. (2017). Liquid droplet formation by HP1 α suggests a role for phase separation in heterochromatin. *Nature* 547, 236–240.
- Lau, M.S., Schwartz, M.G., Kundu, S., Savol, A.J., Wang, P.I., Marr, S.K., Grau, D.J., Schorderet, P., Sadreyev, R.I., Tabin, C.J., et al. (2017). Mutation of a nucleosome compaction region disrupts Polycomb-mediated axial patterning. *Science* (80-.). 355, 1081–1084.
- Lavigne, A.-C., Castells, M., Mermet, J., Kocanova, S., Dalvai, M., and Bystricky, K. (2014). Increased macroH2A1.1 Expression Correlates with Poor Survival of Triple-Negative Breast Cancer Patients. *PLoS One* 9, 1–12.
- Lavigne, M.D., Vatsellas, G., Polyzos, A., Mantouvalou, E., Sianidis, G., Maraziotis, I., Agelopoulos, M., and Thanos, D. (2015). Composite macroH2A/NRF-1 Nucleosomes Suppress Noise and Generate Robustness in Gene Expression. *Cell Rep.* 1–12.
- Lawrence, M., Daujat, S., and Schneider, R. (2016). Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends Genet.* 32, 42–56.
- Leemans, C., Zwalm, M.C.H. Van Der, Brueckner, L., Comoglio, F., Schaik, T. Van, and Pagie, L. (2019). Promoter-Intrinsic and Local Chromatin Features Determine Gene Repression in LADs Article Promoter-Intrinsic and Local Chromatin Features Determine Gene Repression in LADs. *Cell* 177, 852-864.e14.
- de Leeuw, R., Gruenbaum, Y., and Medalia, O. (2018). Nuclear Lamins: Thin Filaments with Major Functions. *Trends Cell Biol.* 28, 34–45.
- Lei, S., Long, J., and Li, J. (2014). MacroH2A suppresses the proliferation of the B16 melanoma cell line. *Mol. Med. Rep.* 10, 1845–1850.
- Li, F., Yi, P., Pi, J., Li, L., Hui, J., and Wang, F. (2016). QKI5-mediated alternative splicing of the histone variant macroH2A1 regulates gastric carcinogenesis. *Oncotarget* 7, 32821–32834.
- Li, X., Kuang, J., Shen, Y., Majer, M.M., Nelson, C.C., Parsawar, K., Heichman, K.A., and Kuwada, S.K. (2012). The atypical histone macroH2A1.2 interacts with HER-2 protein in cancer cells. *J. Biol. Chem.* 287, 23171–23183.
- Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J.P. (2011). Molecular signatures database (MSigDB) 3.0. *27*, 1739–1740.
- Liberzon, A., Birger, C., Thorvaldsdo, H., Ghandi, M., Mesirov, J.P., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database. 417–425.
- Lieberman-aiden, E., Berkum, N.L. Van, Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive Mapping of Long-Range

Interactions Reveals Folding Principles of the Human Genome. 326, 289–294.

López-Terrada, D., Cheung, S.W., Finegold, M.J., and Knowles, B.B. (2009a). Hep G2 is a hepatoblastoma-derived cell line. *Hum. Pathol.* 40, 1512–1515.

López-Terrada, D., Gunaratne, P.H., Adesina, A.M., Pulliam, J., Hoang, D.M., Nguyen, Y., Mistretta, T.A., Margolin, J., and Finegold, M.J. (2009b). Histologic subtypes of hepatoblastoma are characterized by differential canonical Wnt and Notch pathway activation in DLK+ precursors. *Hum. Pathol.* 40, 783–794.

Lopez, R., Sarg, B., Lindner, H., Bartolomé, S., Ponte, I., Suau, P., and Roque, A. (2015). Linker histone partial phosphorylation: Effects on secondary structure and chromatin condensation. *Nucleic Acids Res.* 43, 4463–4476.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. 1–21.

Lu, X., Hamkalo, B., Parseghian, M.H., and Hansen, J.C. (2009). Chromatin Condensing Functions of the Linker Histone C-Terminal Domain Are Mediated by Specific Amino Acid Composition and Intrinsic Protein Disorder. 164–172.

Luger, K. (2006). Dynamic nucleosomes. *Chromosom. Res.* 14, 5–16.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251–260.

Luijsterburg, M.S., de Krijger, I., Wiegant, W.W., Shah, R.G., Smeenk, G., de Groot, A.J.L., Pines, A., Vertegaal, A.C.O., Jacobs, J.J.L., Shah, G.M., et al. (2016). PARP1 Links CHD2-Mediated Chromatin Expansion and H3.3 Deposition to DNA Repair by Non-homologous End-Joining. *Mol. Cell* 61, 547–562.

Lukas, J., Lukas, C., and Bartek, J. (2011). More than just a focus: The chromatin response to DNA damage and its role in genome integrity maintenance. *Nat. Cell Biol.* 13, 1161–1169.

Lupiáñez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R., et al. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* 161, 1012–1025.

Lyko, F. (2018). The DNA methyltransferase family: A versatile toolkit for epigenetic regulation. *Nat. Rev. Genet.* 19, 81–92.

Maeshima, K., Ide, S., and Babokhov, M. (2019). Dynamic chromatin organization without the 30-nm fiber. *Curr. Opin. Cell Biol.* 58, 95–104.

Maiolica, A., de Medina-Redondo, M., Schoof, E.M., Chaikuad, A., Villa, F., Gatti, M., Jeganathan, S., Lou, H.J., Novy, K., Hauri, S., et al. (2014). Modulation of the Chromatin Phosphoproteome by the Haspin Protein Kinase. *Mol. Cell. Proteomics* 13, 1724–1740.

Malladi, S., MacAlinao, D.G., Jin, X., He, L., Basnet, H., Zou, Y., De Stanchina, E., and Massagué, J. (2016). Metastatic Latency and Immune Evasion through Autocrine Inhibition of WNT. *Cell* 165, 45–60.

Mallona, I., Díez-Villanueva, A., Martín, B., and Peinado, M.A. (2017). Chainy: An universal tool for standardized relative quantification in real-time PCR. *Bioinformatics* 33, 1411–1413.

Bibliography

- Mao, B., Wu, W., Li, Y., Hoppe, D., Stannek, P., Glinka, A., and Niehrs, C. (2001). LDL-receptor-related protein 6 is a receptor for Dickkopf proteins. *Nature* **411**, 321–325.
- Margueron, R., Li, G., Sarma, K., Blais, A., Zavadil, J., Woodcock, C.L., Dynlacht, B.D., and Reinberg, D. (2008). Ezh1 and Ezh2 Maintain Repressive Chromatin through Different Mechanisms. *Mol. Cell* **32**, 503–518.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal* **17**, 10.
- Martin, C., Chen, S., Maya-Mendoza, A., Lovric, J., Sims, P.F.G., and Jackson, D. a (2009). Lamin B1 maintains the functional plasticity of nucleoli. *J. Cell Sci.* **122**, 1551–1562.
- Mas, G., Blanco, E., Ballaré, C., Sansó, M., Spill, Y.G., Hu, D., Aoi, Y., Le Dily, F., Shilatifard, A., Marti-Renom, M.A., et al. (2018). Promoter bivalency favors an open chromatin architecture in embryonic stem cells. *Nat. Genet.* **50**, 1452–1462.
- Maze, I., Noh, K.M., Soshnev, A.A., and Allis, C.D. (2014). Every amino acid matters: Essential contributions of histone variants to mammalian development and disease. *Nat. Rev. Genet.* **15**, 259–271.
- Mccoy, A.J. (2007). Solving structures of protein complexes by molecular replacement with Phaser research papers. *Biol. Crystallogr.* **32**–41.
- McGinty, R.K., and Tan, S. (2015). Nucleosome structure and function. *Chem. Rev.* **115**, 2255–2273.
- McStay, B., and Grummt, I. (2008). The epigenetics of rRNA genes: from molecular to chromosome biology. *Annu. Rev. Cell Dev. Biol.* **24**, 131–157.
- Meehan, R.R., Kao, C.F., and Pennings, S. (2003). HP1 binding to native chromatin in vitro is determined by the hinge region and not by the chromodomain. *EMBO J.* **22**, 3164–3174.
- Mermoud, J.E., Costanzi, C., Pehrson, J.R., and Brockdorff, N. (1999). Histone macroH2A1.2 relocates to the inactive X chromosome after initiation and propagation of X-inactivation. *J. Cell Biol.* **147**, 1399–1408.
- Messner, S., and Hottiger, M.O. (2011). Histone ADP-ribosylation in DNA repair, replication and transcription. *Trends Cell Biol.* **21**, 534–542.
- Meuleman, W., Peric-hupkes, D., Kind, J., Beaudry, J., Pagie, L., Kellis, M., Reinders, M., Wessels, L., and Steensel, B. Van (2013). Constitutive nuclear lamina – genome interactions are highly conserved and associated with A / T-rich sequence. *Genome Res.* **23**, 270–280.
- Miga, K.H. (2015). Completing the human genome: the progress and challenge of satellite DNA assembly. *Chromosom. Res.* **23**, 421–426.
- Mootha, V.K., Lindgren, C.M., Eriksson, K., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E., et al. (2003). PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Proc. Natl. Acad. Sci. USA* **100**, 267–273.
- Moreland, J.L., Gramada, A., Buzko, O. V, Zhang, Q., and Bourne, P.E. (2005). The Molecular Biology Toolkit (MBT): a modular platform for developing molecular visualization applications. *BMC Bioinformatics* **6**, 21.

- Morgenstern, J.P., and Land, H. (1990). Advanced mammalian gene transfer: High titre retroviral vectors with multiple drug selection markers and a complementary helper-free packaging cell line. *Nucleic Acids Res.* 18, 3587–3596.
- Musselman, C.A., Lalonde, M.E., Côté, J., and Kutateladze, T.G. (2012). Perceiving the epigenetic landscape through histone readers. *Nat. Struct. Mol. Biol.* 19, 1218–1227.
- Muthurajan, U.M., McBryant, S.J., Lu, X., Hansen, J.C., and Luger, K. (2011). The linker region of macroH2A promotes self-association of nucleosomal arrays. *J. Biol. Chem.* 286, 23852–23864.
- Nashun, B., Yukawa, M., Liu, H., Akiyama, T., and Aoki, F. (2010). Changes in the nuclear deposition of histone H2A variants during pre-implantation development in mice. *Development* 137, 3785–3794.
- Németh, A., Conesa, A., Santoyo-Lopez, J., Medina, I., Montaner, D., Péterfia, B., Solovei, I., Cremer, T., Dopazo, J., and Längst, G. (2010a). Initial genomics of the human nucleolus. *PLoS Genet.* 6, e1000889.
- Németh, A., Conesa, A., Santoyo-Lopez, J., Medina, I., Montaner, D., Péterfia, B., Solovei, I., Cremer, T., Dopazo, J., and Längst, G. (2010b). Initial Genomics of the Human Nucleolus. *PLoS Genet.* 6, e1000889.
- Neri, F., Rapelli, S., Krepelova, A., Incarnato, D., Parlato, C., Basile, G., Maldotti, M., Anselmi, F., and Oliviero, S. (2017). Intragenic DNA methylation prevents spurious transcription initiation. *Nature* 543, 72–77.
- Nora, E.P., Goloborodko, A., Valton, A.-L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* 169, 930-944.e22.
- Novikov, L., Park, J.W., Chen, H., Klerman, H., Jalloh, A.S., and Gamble, M.J. (2011). QKI-Mediated Alternative Splicing of the Histone Variant MacroH2A1 Regulates Cancer Cell Proliferation. *Mol. Cell. Biol.* 31, 4244–4255.
- Nuebler, J., Fudenberg, G., Imakaev, M., Abdennur, N., and Mirny, L.A. (2018). Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc. Natl. Acad. Sci. U. S. A.* 115, E6697–E6706.
- Nusinow, D.A., Sharp, J.A., Morris, A., Salas, S., Plath, K., and Panning, B. (2007a). The Histone Domain of macroH2A1 Contains Several Dispersed Elements that Are Each Sufficient to Direct Enrichment on the Inactive X Chromosome. *J. Mol. Biol.* 371, 11–18.
- Nusinow, D.A., Hernández-Muñoz, I., Fazzio, T.G., Shah, G.M., Kraus, W.L., and Panning, B. (2007b). Poly (ADP-ribose) Polymerase 1 Is Inhibited by a Histone H2A Variant, MacroH2A, and Contributes to Silencing of the Inactive X Chromosome * □ and Barbara Panning. 282, 12851–12859.
- O'Carroll, D., Scherthan, H., Peters, A.H.F.M., Opravil, S., Haynes, A.R., Laible, G., Rea, S., Schmid, M., Lebersorger, A., Jerratsch, M., et al. (2000). Isolation and Characterization of Suv39h2, a Second Histone H3 Methyltransferase Gene That Displays Testis-Specific Expression. *Mol. Cell. Biol.* 20, 9423–9433.
- Ogiyama, Y., Schuettengruber, B., Papadopoulos, G.L., Chang, J.M., and Cavalli, G. (2018). Polycomb-Dependent Chromatin Looping Contributes to Gene Silencing during Drosophila Development. *Mol. Cell* 71, 73-88.e5.

Bibliography

- Olins, A.L., Olins, D.E., Science, S., Series, N., and Jan, N. (1974). Spheroid Chromatin Units (v Bodies). *183*, 330–332.
- Ou, H.D., Phan, S., Deerinck, T.J., Thor, A., Ellisman, M.H., and O'Shea, C.C. (2017). ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. *Science* (80-.). 357.
- Ouararhni, K., Hadj-Slimane, R., Ait-Si-Ali, S., Robin, P., Mietton, F., Harel-Bellan, A., Dimitrov, S., and Hamiche, A. (2006). The histone variant mH2A1.1 interferes with transcription by down-regulating PARP-1 enzymatic activity. *Genes Dev.* 20, 3324–3336.
- Padeken, J., and Heun, P. (2014). Nucleolus and nuclear periphery: Velcro for heterochromatin. *Curr. Opin. Cell Biol.* 28, 54–60.
- Padeken, J., Zeller, P., and Gasser, S.M. (2015). Repeat DNA in genome organization and stability. *Curr. Opin. Genet. Dev.* 31, 12–19.
- Pandur, P., Läsche, M., Eisenberg, L.M., and Kühl, M. (2002). Wnt-11 activation of a non-canonical Wnt signalling pathway is required for cardiogenesis. *Nature* 418, 636–641.
- Park, I.K., Qian, D., Kiel, M., Becker, M.W., Pihalja, M., Weissman, I.L., Morrison, S.J., and Clarke, M.F. (2003). Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. *Nature* 423, 302–305.
- Park, S.-J., Shim, J.W., Park, H.S., Eum, D.-Y., Park, M.-T., Mi Yi, J., Choi, S.H., Kim, S.D., Son, T.G., Lu, W., et al. (2016). MacroH2A1 downregulation enhances the stem-like properties of bladder cancer cells by transactivation of Lin28B. *Oncogene* 35, 1292–1301.
- Pascual-Reguant, L., Blanco, E., Galan, S., Le Dily, F., Cuartero, Y., Serra-Bardenys, G., Di Carlo, V., Iturbide, A., Cebrià-Costa, J.P., Nonell, L., et al. (2018). Lamin B1 mapping reveals the existence of dynamic and functional euchromatin lamin B1 domains. *Nat. Commun.* 9.
- Pasque, V., Gillich, A., Garrett, N., and Gurdon, J.B. (2011a). Histone variant macroH2A confers resistance to nuclear reprogramming. *EMBO J.* 30, 2373–2387.
- Pasque, V., Halley-stott, R.P., Gillich, A., Garrett, N., and John, B. (2011b). Epigenetic stability of repressed states involving the histone variant macroH2A revealed by nuclear transfer to *Xenopus* oocytes. *Nucleus* 2, 533–539.
- Pasque, V., Radzisheuskaya, a., Gillich, a., Halley-Stott, R.P., Panamarova, M., Zernicka-Goetz, M., Surani, M. a., and Silva, J.C.R. (2012). Histone variant macroH2A marks embryonic differentiation in vivo and acts as an epigenetic barrier to induced pluripotency. *J. Cell Sci.* 2, 6094–6104.
- Passarge, E. (1979). Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago. *Am. J. Hum. Genet.* 31, 106.
- Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 14, 417–419.
- Pehrson, J.R., and Fried, V. (1992). MacroH2A, a core histone containing a large nonhistone region. *Science* 257, 1398–1400.

- Pehrson, J.R., Costanzi, C., and Dharia, C. (1997). Developmental and tissue expression patterns of histone macroH2A1 subtypes. *J. Cell. Biochem.* 65, 107–113.
- Pehrson, J.R., Changolkar, L.N., Costanzi, C., and Leu, N.A. (2014). Mice Without MacroH2A Histone Variants. *Mol. Cell. Biol.* 34, 4523–4533.
- Pękowska, A., Klaus, B., Xiang, W., Severino, J., Daigle, N., Klein, F.A., Oleś, M., Casellas, R., Ellenberg, J., Steinmetz, L.M., et al. (2018). Gain of CTCF-Anchored Chromatin Loops Marks the Exit from Naive Pluripotency. *Cell Syst.* 7, 482–495.e10.
- Pengelly, A.R., Kalb, R., Finkl, K., and Müller, J. (2015). Transcriptional repression by PRC1 in the absence of H2A monoubiquitylation. *Genes Dev.* 29, 1487–1492.
- Peric-Hupkes, D., Meuleman, W., Pagie, L., Bruggeman, S.W.M., Solovei, I., Brugman, W., Gräf, S., Flicek, P., Kerkhoven, R.M., van Lohuizen, M., et al. (2010). Molecular Maps of the Reorganization of Genome-Nuclear Lamina Interactions during Differentiation. *Mol. Cell* 38, 603–613.
- Perkins, J.R., Dawes, J.M., McMahon, S.B., Bennett, D.L.H., Orenge, C., and Kohl, M. (2012). ReadqPCR and NormqPCR: R packages for the reading, quality checking and normalisation of RT-qPCR quantification cycle (Cq) data. *BMC Genomics* 13.
- Peters, A.H.F.M., Carroll, O., Scherthan, H., Mechtler, K., Sauer, S., Scho, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A., et al. (2001). Loss of the Suv39h Histone Methyltransferases Impairs Mammalian Heterochromatin and Genome Stability. *Cell* 107, 323–337.
- Phillips-Cremins, J.E., Sauria, M.E.G., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S.K., Ong, C.T., Hookway, T.A., Guo, C., Sun, Y., et al. (2013). Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* 153, 1281–1295.
- Pliatska, M., Kapasa, M., Kokkalis, A., Polyzos, A., and Thanos, D. (2018). The histone variant macroH2A blocks cellular reprogramming by inhibiting mesenchymal-to-epithelial transition. *Mol. Cell. Biol.* MCB.00669-17.
- Pogo, B.G., Allfrey, V.G., and Mirsky, A.E. (1966). RNA synthesis and histone acetylation during the course of gene activation in lymphocytes. *Proc. Natl. Acad. Sci. U. S. A.* 55, 805–812.
- Polioudaki, H., Kourmouli, N., Drosou, V., Bakou, A., Theodoropoulos, P.A., Singh, P.B., Giannakouros, T., and Georgatos, S.D. (2001). Histones H3/H4 form a tight complex with the inner nuclear membrane protein LBR and heterochromatin protein 1. *EMBO Rep.* 2, 920–925.
- Posavec-Marjanović, M., Hurtado-Bagès, S., Lassi, M., Valero, V., Malinverni, R., Delage, H., Navarro, M., Corujo, D., Guberovic, I., Douet, J., et al. (2017). MacroH2A1 .1 regulates mitochondrial respiration by limiting nuclear NAD⁺ consumption. *Nat. Struct. Mol. Biol.* 24, 902–910.
- Posavec, M., Timinszky, G., and Buschbeck, M. (2013). Macro domains as metabolite sensors on chromatin. *Cell. Mol. Life Sci.* 70, 1509–1524.
- Quinodoz, S.A., Ollikainen, N., Tabak, B., Palla, A., Schmidt, J.M., Detmar, E., Lai, M.M., Shishkin, A.A., Bhat, P., Takei, Y., et al. (2018). Higher-Order Inter-chromosomal Hubs Shape 3D Genome Organization in the Nucleus. *Cell* 1–14.

Bibliography

- Rao, S.S.P., Huntley, M.H., Durand, N.C., and Stamenova, E.K. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* 159, 1665–1680.
- Rasmussen, K.D., and Helin, K. (2016). Role of TET enzymes in DNA methylation, development, and cancer. *Genes Dev.* 30, 733–750.
- Lo Re, O., Fusilli, C., Rappa, F., Van Haele, M., Douet, J., Pindjakova, J., Wanessa Rocha, S., Pata, I., Valcikova, B., Uldrijan, S., et al. (2017). Induction of Cancer Cell Stemness by Depletion of Macrohistone H2A1 in Hepatocellular Carcinoma. *Hepatology*.
- Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtier, K., Ponting, C.P., Allis, C.D., et al. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406, 593–599.
- Reddy, K.L., Zullo, J.M., Bertolino, E., and Singh, H. (2008). Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* 452, 243–247.
- Ricci, M.A., Manzo, C., García-Parajo, M.F., Lakadamyali, M., and Cosma, M.P. (2015). Chromatin Fibers Are Formed by Heterogeneous Groups of Nucleosomes In Vivo. *Cell* 160, 1145–1158.
- Riggs, A.D. (1975). X inactivation, differentiation, and DNA methylation. *Cytogenet. Cell Genet.* 14, 9–25.
- Ritz, C., and Spiess, A.N. (2008). qpcR: An R package for sigmoidal model selection in quantitative real-time polymerase chain reaction analysis. *Bioinformatics* 24, 1549–1551.
- Rivera-Casas, C., Gonzalez-Romero, R., Cheema, M.S., Ausió, J., and Eirín-López, J.M. (2016). The characterization of macroH2A beyond vertebrates supports an ancestral origin and conserved role for histone variants in chromatin. *Epigenetics* 11, 415–425.
- Robson, M.I., de las Heras, J.I., Czapiewski, R., Lê Thành, P., Booth, D.G., Kelly, D.A., Webb, S., Kerr, A.R.W., and Schirmer, E.C. (2016). Tissue-Specific Gene Repositioning by Muscle Nuclear Membrane Proteins Enhances Repression of Critical Developmental Genes during Myogenesis. *Mol. Cell* 62.
- Rose, A.S., and Hildebrand, P.W. (2015). NGL Viewer: a web application for molecular visualization. *Nucleic Acids Res.* 43, W576–W579.
- Rowley, M.J., Nichols, M.H., Lyu, X., Ando-Kuri, M., Rivera, I.S.M., Hermetz, K., Wang, P., Ruan, Y., and Corces, V.G. (2017). Evolutionarily Conserved Principles Predict 3D Chromatin Organization. *Mol. Cell* 67, 837–852.e7.
- Rubin, A.J., Barajas, B.C., Furlan-Magaril, M., Lopez-Pajares, V., Mumbach, M.R., Howard, I., Kim, D.S., Boxer, L.D., Cairns, J., Spivakov, M., et al. (2017). Lineage-specific dynamic and pre-established enhancer-promoter contacts cooperate in terminal differentiation. *Nat. Genet.* 49, 1522–1528.
- Ruiz, P.D., and Gamble, M.J. (2018). MacroH2A1 chromatin specification requires its docking domain and acetylation of H2B lysine 20. *Nat. Commun.* 9.
- Sakabe, T., Azumi, J., Umekita, Y., Toriguchi, K., Hatano, E., Hirooka, Y., and Shiota, G. (2017). Expression of Cancer Stem Cell-associated DKK1 mRNA Serves as Prognostic Marker for Hepatocellular Carcinoma. *Anticancer Res.* 37, 4881–4888.

- Saksouk, N., Barth, T.K., Ziegler-Birling, C., Olova, N., Nowak, A., Rey, E., Mateos-Langerak, J., Urbach, S., Reik, W., Torres-Padilla, M.-E., et al. (2014). Redundant Mechanisms to Form Silent Chromatin at Pericentromeric Regions Rely on BEND3 and DNA Methylation. *Mol. Cell* 56, 580–594.
- Saksouk, N., Simboeck, E., and Dejardin, J. (2015). Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin* 8, 3.
- Salama, R., Sadaie, M., Hoare, M., and Narita, M. (2014). Cellular senescence and its effector programs. *Genes Dev.* 28, 99–114.
- Sanborn, A.L., Rao, S.S.P., Huang, S.C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J., et al. (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci. U. S. A.* 112, E6456–E6465.
- Sati, S., and Cavalli, G. (2017). Chromosome conformation capture technologies and their impact in understanding genome function. *Chromosoma* 126, 33–44.
- Sato, H., Wu, B., Delahaye, F., Singer, R.H., and Grealley, J.M. (2019). Retargeting of macroH2A following mitosis to cytogenetic-scale heterochromatic domains. *J. Cell Biol.* jcb.201811109.
- Sato, N., Yamabuki, T., Takano, A., Koinuma, J., Aragaki, M., Masuda, K., Ishikawa, N., Kohno, N., Ito, H., Miyamoto, M., et al. (2010). Wnt inhibitor Dickkopf-1 as a target for passive cancer immunotherapy. *Cancer Res.* 70, 5326–5336.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: An open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682.
- Schnater, J.M., Köhler, S.E., Lamers, W.H., Von Schweinitz, D., and Aronson, D.C. (2003). Where do we stand with hepatoblastoma? A review. *Cancer* 98, 668–678.
- Schoenfelder, S., Furlan-magaril, M., Mifsud, B., Tavares-cadete, F., Sugar, R., Javierre, B., Nagano, T., Katsman, Y., Sakthidevi, M., Wingett, S.W., et al. (2015). The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements. *Genome Res.* 582–597.
- Schuettengruber, B., Bourbon, H., Croce, L. Di, and Cavalli, G. (2017). Genome Regulation by Polycomb and Trithorax : 70 Years and Counting. *Cell* 171, 34–57.
- Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., and Rauscher, F.J. (2002). SETDB1: A novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev.* 16, 919–932.
- Schwartzman, O., Mukamel, Z., Oded-Elkayam, N., Olivares-Chauvet, P., Lubling, Y., Landan, G., Izraeli, S., and Tanay, A. (2016). UMI-4C for quantitative and targeted chromosomal contact profiling. *Nat. Methods* 13, 685–691.
- Sellou, H., Lebeaupin, T., Chapuis, C., Smith, R., Hegele, A., and Tansey, W.P. (2016). The poly (ADP-ribose) -dependent chromatin remodeler Alc1 induces local chromatin relaxation upon DNA damage. 27.

Bibliography

- Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148**, 458–472.
- Shah, P.P., Donahue, G., Otte, G.L., Capell, B.C., Nelson, D.M., Cao, K., Aggarwala, V., Cruickshanks, H.A., Rai, T.S., McBryan, T., et al. (2013). Lamin B1 depletion in senescent cells triggers large-scale changes in gene expression and the chromatin landscape. *Genes Dev.* **27**, 1787–1799.
- Sharma, A.B., Dimitrov, S., Hamiche, A., and Van Dyck, E. (2019). Centromeric and ectopic assembly of CENP-A chromatin in health and cancer: old marks and new tracks. *Nucleic Acids Res.* **47**, 1051–1069.
- Sheedfar, F., Vermeer, M., Paziienza, V., Villarroja, J., Rappa, F., Cappello, F., Mazzocchi, G., Villarroja, F., van der Molen, H., Hofker, M.H., et al. (2015). Genetic ablation of macrohistone H2A1 leads to increased leanness, glucose tolerance and energy expenditure in mice fed a high-fat diet. *Int. J. Obes. (Lond)*. **39**, 331–338.
- Shen, H., and Laird, P.W. (2013). Interplay between the cancer genome and epigenome. *Cell* **153**, 38–55.
- Shen, L., Shao, N., Liu, X., and Nestler, E. (2014). Ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics* **15**.
- Shen, Q., Fan, J., Yang, X.R., Tan, Y., Zhao, W., Xu, Y., Wang, N., Niu, Y., Wu, Z., Zhou, J., et al. (2012a). Serum DKK1 as a protein biomarker for the diagnosis of hepatocellular carcinoma: A large-scale, multicentre study. *Lancet Oncol.* **13**, 817–826.
- Shen, X., Liu, Y., Hsu, Y.J., Fujiwara, Y., Kim, J., Mao, X., Yuan, G.C., and Orkin, S.H. (2008). EZH1 Mediates Methylation on Histone H3 Lysine 27 and Complements EZH2 in Maintaining Stem Cell Identity and Executing Pluripotency. *Mol. Cell* **32**, 491–502.
- Shen, Y., Yue, F., Mc Cleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Ren, B., et al. (2012b). A map of the cis-regulatory sequences in the mouse genome. *Nature* **488**, 116–120.
- Shevelyov, Y.Y., and Ulianov, S. V. (2019). The Nuclear Lamina as an Organizer of Chromosome Architecture. *Cells* **8**, 136.
- Shibata, S., Asano, T., Naito, A.N.M., Ogurat, A., and Division, K.D.O.I. (1998). Combined Immunodeficient Mice Transplanted With Human Peripheral Blood Lymphocytes. *Immunology* **93**, 524–532.
- Shinagawa, T., Takagi, T., Tsukamoto, D., Tomaru, C., Huynh, L.M., Sivaraman, P., Kumarevel, T., Inoue, K., Nakato, R., Katou, Y., et al. (2014). Histone Variants Enriched in Oocytes Enhance Reprogramming to Induced Pluripotent Stem Cells. *Cell Stem Cell* **14**, 217–227.
- Shlyueva, D., Stampfel, G., and Stark, A. (2014). Transcriptional enhancers: From properties to genome-wide predictions. *Nat. Rev. Genet.* **15**, 272–286.
- Simon, J.M., Giresi, P.G., Davis, I.J., and Lieb, J.D. (2012). Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA. *Nat. Protoc.* **7**, 256–267.
- Slee, R.B., Steiner, C.M., Herbert, B.S., Vance, G.H., Hickey, R.J., Schwarz, T., Christan, S., Radovich,

- M., Schneider, B.P., Schindelhauer, D., et al. (2012). Cancer-associated alteration of pericentromeric heterochromatin may contribute to chromosome instability. *Oncogene* 31, 3244–3253.
- Smith, Z.D., and Meissner, A. (2013). DNA methylation: Roles in mammalian development. *Nat. Rev. Genet.* 14, 204–220.
- Söderberg, O., Gullberg, M., Jarvius, M., Ridderstrale, K., Leuchowius, K.-J., Jarvius, J., Wester, K., Hydbring, P., Bahram, F., Larsson, L.-G., et al. (2006). Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nat Methods* 3, 995–1000.
- Soneson, C., Love, M.I., and Robinson, M.D. (2016). Differential analyses for RNA-seq: Transcript-level estimates improve gene-level inferences [version 2; referees: 2 approved]. *F1000Research* 4, 1–22.
- Soufi, A., Donahue, G., and Zaret, K.S. (2012). Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* 151, 994–1004.
- Spector, L.G., and Birch, J. (2012). The epidemiology of hepatoblasoma. *Pediatr. Blood Cancer* 59, 779.
- Sporn, J.C., and Jung, B. (2012). Differential Regulation and Predictive Potential of MacroH2A1 Isoforms in Colon Cancer. *AJPA* 180, 2516–2526.
- Sporn, J.C., Kustatscher, G., Hothorn, T., Collado, M., Serrano, M., Muley, T., Schnabel, P., and Ladurner, A.G. (2009). Histone macroH2A isoforms predict the risk of lung cancer recurrence. *Oncogene* 28, 3423–3428.
- Stadhouders, R., Vidal, E., Serra, F., Di Stefano, B., Le Dily, F., Quilez, J., Gomez, A., Collombet, S., Berenguer, C., Cuartero, Y., et al. (2018). Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. *Nat. Genet.* 50, 238–249.
- Stadhouders, R., Filion, G.J., and Graf, T. (2019). Transcription factors and 3D genome conformation in cell-fate decisions. *Nature* 569, 345–354.
- van Steensel, B., and Belmont, A.S. (2017). Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell* 169, 780–791.
- Stephens, A.D., Liu, P.Z., Banigan, E.J., Almassalha, L.M., Backman, V., Adam, S.A., Goldman, R.D., and Marko, J.F. (2018). Chromatin histone modifications and rigidity affect nuclear morphology independent of lamins. *Mol. Biol. Cell* 29, 220–233.
- Stephens, A.D., Banigan, E.J., and Marko, J.F. (2019). Chromatin's physical properties shape the nucleus and its functions. *Curr. Opin. Cell Biol.* 58, 76–84.
- van der Stoop, P., Boutsma, E.A., Hulsman, D., Noback, S., Heimerikx, M., Kerkhoven, R.M., Voncken, J.W., Wessels, L.F.A., and van Lohuizen, M. (2008). Ubiquitin E3 ligase Ring1b/Rnf2 of polycomb repressive complex 1 contributes to stable maintenance of mouse embryonic stem cells. *PLoS One* 3, e2235.
- Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. *Nature* 403, 41–45.
- Strickfaden, H., McDonald, D., Kruhlak, M.J., Haince, J., Th, J.P.H., Rouleau, M., Ishibashi, T., Corry, G.N., Ausio, J., Underhill, D.A., et al. (2016). Poly (ADP-ribosyl) ation-dependent Transient Chromatin

Bibliography

Decondensation and Histone Displacement following Laser. *291*, 1789–1802.

Strom, A.R., Alexander, V., Emelyanov, V., Mir, M., Fyodorov, D. V., Darzacq, X., and Karpen, G.H. (2017). Phase separation drives heterochromatin domain formation. *Nat. Publ. Gr.*

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., and Ebert, B.L. (2005). Gene set enrichment analysis : A knowledge-based approach for interpreting genome-wide.

Sun, Z., Filipescu, D., Andrade, J., Gaspar-Maia, A., Ueberheide, B., and Bernstein, E. (2018). Transcription-associated histone pruning demarcates macroH2A chromatin domains. *Nat. Struct. Mol. Biol.* *25*, 958–970.

Szabo, Q., Bantignies, F., and Cavalli, G. (2019). Principles of genome folding into topologically associating domains. *Sci. Adv.* *5*.

Szczepińska, T., Rusek, A.M., and Plewczynski, D. (2019). Intermingling of chromosome territories. *Genes Chromosom. Cancer* *58*, 500–506.

Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H., et al. (2002). G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev.* *16*, 1779–1791.

Tachibana, M., Ueda, J., Fukuda, M., Takeda, N., Ohta, T., Iwanari, H., Sakihama, T., Kodama, T., Hamakubo, T., and Shinkai, Y. (2005). Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes Dev.* *19*, 815–826.

Taddei, A., Maison, C., Roche, D., and Almouzni, G. (2001). Reversible disruption of pericentric heterochromatin and centromere function by inhibiting deacetylases. *Nat. Cell Biol.* *3*, 114–120.

Talbert, P.B., and Henikoff, S. (2010). Histone variants — ancient wrap artists of the epigenome. *Nat. Rev. Mol. Cell Biol.* *11*, 264–275.

Tanaka, Y., Inoue, T., and Horie, H. (2013). International pediatric liver cancer pathological classification: Current trend. *Int. J. Clin. Oncol.* *18*, 946–954.

Tanasijevic, B., and Rasmussen, T.P. (2011). X Chromosome inactivation and differentiation occur readily in ES cells doubly-deficient for macroH2A1 and macroH2A2. *PLoS One* *6*.

Tao, Y.M., Liu, Z., and Liu, H.L. (2013). Dickkopf-1 (DKK1) promotes invasion and metastasis of hepatocellular carcinoma. *Dig. Liver Dis.* *45*, 251–257.

Tavares, L., Dimitrova, E., Oxley, D., Webster, J., Poot, R., Demmers, J., Bezstarosti, K., Taylor, S., Ura, H., Koide, H., et al. (2012). RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. *Cell* *148*, 664–678.

Thiry, M., Lepoint, A., and Goessens, G. (1985). Re-evaluation of the site of transcription in Ehrlich tumour cell nucleoli. *Biol. Cell* *54*, 57–64.

Timinszky, G., Till, S., Hassa, P.O., Hothorn, M., Kustatscher, G., Nijmeijer, B., Colombelli, J., Altmeyer, M., Stelzer, E.H.K., Scheffzek, K., et al. (2009). A macrodomain-containing histone rearranges chromatin upon sensing PARP1 activation. *Nat. Struct. Mol. Biol.* *16*, 923–929.

Ting, D.T., Ting, D.T., Lipson, D., Paul, S., Brannigan, B.W., Akhavanfard, S., Coffman, E.J., Contino, G., Deshpande, V., Iafrate, A.J., et al. (2011). Aberrant Overexpression of Satellite Epithelial Cancers. 593, 593–597.

Towbin, B.D., González-Aguilera, C., Sack, R., Gaidatzis, D., Kalck, V., Meister, P., Askjaer, P., and Gasser, S.M. (2012). Step-Wise Methylation of Histone H3K9 Positions Heterochromatin at the Nuclear Periphery. *Cell* 150, 934–947.

Treier, M., Gleiberman, A.S., Connell, S.M.O., Szeto, D.P., McMahon, J.A., McMahon, A.P., and Rosenfeld, M.G. (1998). Multistep signaling requirements for pituitary organogenesis in vivo. 6, 1691–1704.

Ulianov, S. V., Khrameeva, E.E., Gavrillov, A.A., Flyamer, I.M., Kos, P., Mikhaleva, E.A., Penin, A.A., Logacheva, M.D., Imakaev, M. V., Chertovich, A., et al. (2016). Active chromatin and transcription play a key role in chromosome partitioning into topologically associating domains. *Genome Res.* 26, 70–84.

Vicent, G.P., Nacht, A.S., Font-Mateu, J., Castellano, G., Gaveglia, L., Ballaré, C., and Beato, M. (2011). Four enzymes cooperate to displace histone H1 during the first minute of hormonal gene activation. *Genes Dev.* 25, 845–862.

Vieira-Silva, T.S., Monteiro-Reis, S., Barros-Silva, D., Ramalho-Carvalho, J., Graça, I., Carneiro, I., Martins, A.T., Oliveira, J., Antunes, L., Hurtado-Bagès, S., et al. (2019). Histone variant MacroH2A1 is downregulated in prostate cancer and influences malignant cell phenotype. *Cancer Cell Int.* 19, 1–13.

Voon, H.P.J., and Wong, L.H. (2016). New players in heterochromatin silencing : histone variant H3 . 3 and the ATRX / DAXX chaperone. 1–6.

Waddington, C.H. (1942). The epigenotype. *Endeavour* 18–20.

Waddington, C.H. (1957). The strategy of the genes. A discussion of some aspects of theoretical biology. (Allen & Unwin).

Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., and Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. *Nature* 431, 873–878.

Wang, S., Su, J., Beliveau, B.J., Bintu, B., Moffitt, J.R., and Wu, C. (2016). Spatial organization of chromatin domains and compartments in single chromosomes. *Science* (80-.). 353, 598–602.

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., and Barton, G.J. (2009). Jalview Version 2-A multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191.

Wen, B., Wu, H., Shinkai, Y., Irizarry, R.A., and Feinberg, A.P. (2009). Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat. Genet.* 41, 246–250.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis Using the Grammar of Graphics* (Springer-Verlag New York).

Wu, B., Davey, G.E., Nazarov, A.A., Dyson, P.J., and Davey, C.A. (2011). Specific DNA structural attributes modulate platinum anticancer drug site selection and cross-link generation. *Nucleic Acids Res.*

Bibliography

39, 8200–8212.

Xie, H., Xu, J., Hsu, J.H., Nguyen, M., Fujiwara, Y., Peng, C., and Orkin, S.H. (2014). Polycomb repressive complex 2 regulates normal hematopoietic stem cell function in a developmental-stage-specific manner. *Cell Stem Cell* 14, 68–80.

Xu, J., and Liu, Y. (2019). A guide to visualizing the spatial epigenome with super-resolution microscopy. *FEBS J.* 286, 3095–3109.

Yang, H., Chen, G.-D., Fang, F., Liu, Z., Hiu Yan Lau, S., Zhang, J.-F., Lau, W.Y., and Yang, L.-Y. (2013). Dickkopf-1: As a Diagnostic and Prognostic Serum Marker for Early Hepatocellular Carcinoma. *Int. J. Biol. Markers* 28, 286–297.

Ye, Q., Callebaut, I., Pezhman, A., Courvalin, J.C., and Worman, H.J. (1997). Domain-specific interactions of human HP1-type chromodomain proteins and inner nuclear membrane protein LBR. *J. Biol. Chem.* 272, 14983–14989.

Yelagandula, R., Stroud, H., Holec, S., Zhou, K., Feng, S., Zhong, X., Muthurajan, U.M., Nie, X., Kawashima, T., Groth, M., et al. (2014). The Histone Variant H2A.W Defines Heterochromatin and Promotes Chromatin Condensation in Arabidopsis. *Cell* 158, 98–109.

Zhang, J., Zhao, Y., and Yang, Q. (2014). Sensitivity and Specificity of Dickkopf-1 Protein in Serum for Diagnosing Hepatocellular Carcinoma: A Meta-Analysis. *Int. J. Biol. Markers* 29, 403–410.

Zhang, R., Poustovoitov, M. V., Ye, X., Santos, H.A., Chen, W., Daganzo, S.M., Erzberger, J.P., Serebriiskii, I.G., Canutescu, A.A., Dunbrack, R.L., et al. (2005a). Formation of MacroH2A-Containing Senescence-Associated Heterochromatin Foci and Senescence Driven by ASF1a and HIRA. *Dev. Cell* 8, 19–30.

Zhang, R., Poustovoitov, M. V., Ye, X., Santos, H.A., Chen, W., Daganzo, S.M., Erzberger, J.P., Serebriiskii, I.G., Canutescu, A.A., Dunbrack, R.L., et al. (2005b). Formation of MacroH2A-Containing Senescence-Associated Heterochromatin Foci and Senescence Driven by ASF1a and HIRA. *Dev. Cell* 8, 19–30.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 9, R137.

Zhao, Y., and Garcia, B.A. (2015). Comprehensive catalog of currently documented histone modifications. *Cold Spring Harb. Perspect. Biol.* 7, 1–21.

Zillner, K., Komatsu, J., Filarsky, K., Kalepu, R., Bensimon, A., and Nemeth, A. (2015). Active human nucleolar organizer regions are interspersed with inactive rDNA repeats in normal and tumor cells. *Epigenomics* 7, 363–378.

Zimmermann, A. (2005). The emerging family of hepatoblastoma tumours: From ontogenesis to oncogenesis. *Eur. J. Cancer* 41, 1503–1514.

